

FIGURE 1

ACTGCACCTCGGTTCTATCGATTGAATTCCTCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA
CCCACGCGTCCGGGCGGAGCAGCACGCGCCGAGGACCTGGAGTCCGGCTCGCTCTTCCCG
CAGCGCTACCCGCCATGCGCCTGCCGCGCCGGGCCGCTGGGGCTCCTGCCGCTCTTGCTG
CTGCTGCCGCCCGCGCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCT
GGTGGACAAGTTTAACGAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGCGGGGAACA
CGGCTTGGGAGGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTTCGCTGCTGGAGATC
CTGGAGGGGCTGTGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGA
GCACCTGGAGGCTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTT
GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCCGACTGTCTCGCATGC
CAGGGCGGATCCAGAGGCCCTGCAGCGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA
GGCGCAGGGTCTCTGCCGTGCCACATGGGGTACCAGGGCCCGCTGTGCACTGACTGCATGG
ACGGCTACTTTCAGCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC
TGCAAGACGTGCTCGGGCTGACCAACAGAGACTGCGCGAGTGTGAAGTGGGCTGGGTGCT
GGACGAGGGCGCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCTCCCTGCAGCGCTG
CGCAGTTCTGTAAGAACGCCAACGGCTCCTACCGTGCAGAGAGTGTGACTCCAGCTGTGTG
GGCTGCACAGGGGAAGGCCAGGAACTGTAAGAGTGTATCTCTGGCTACGCGAGGGAGCA
CGGACAGTGTGCAGATGTGGACGAGTGTCTACTAGCAGAAAAAACCTGTGTGAGGAAAAACG
AAAACTGCTACAATACTCCAGGAGCTACGTCTGTGTGTCTGACGGCTTCGAAGAAACG
GAAGATGCCCTGTGTGCCCGCGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCT
GCCCTCCCGCGAAGACCTGTAAATGTGCCGGACTTACCCTTTAAATTATTAGAAGGATGTCC
CGTGGAAAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGACAGCGCGGGGAGAGGCTGC
CTGCTCTCTAACGGTTGATTCTCATTGTGCCCTTAAACAGCTGCATTCTTGGTTGTTCTTA
AACAGACTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG
AGGAAAAAAAAAAAAAAAAAAAAAGGCGCGCGCACTCTAGAGTCGACCTGCAGAAGC
TTGGCCGCCATGGCCCACTTGTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA
TCACAAATTTACAAATAAAGCATTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACT
ATCAATGTATCTTATCATGTCTGGATCGGGAATTAATTCGGCGCAGCACCATTGGCCTGAAAT
AACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG
TGTGTCAATTAGGGTGTGGAAGTCCCAGGCTCCCAGCAGGCAGAAGTATGCAAGCATGC
ATCTCAATTAGTCAGCAACCCAGTTTT

FIGURE 2

><subunit 1 of 1, 353 aa, 0 stop

><MW: 38192, pI: 4.53, NX(S/T): 2

MRLPRRAALGLLPLLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEKTL SKYESSEIRL
LEILEGLCESSDFECNQMLEAQEBEHLAWWLQLKSEYPDLFEWFVCVTKLVCCSPGTYGPDCLACQGGSGRPCSG
NGHCSGDGSRQGDGSCRCMHGYQGPLCTDCMDGYFSSLRNETHS ICTACDESKTC SGLTNRDGCEVGVWLDE
GACVDVDECAAEPPPCSAAQFCKNANGSYTCECDSSCVGCTGEGPGNCKECISGVAREHGQCADVDECSLAEKT
CVRKNENCYNTPGSYVVCVCPDGFEEETEDACVPPAEAEATEGESPTQLPSREDL

Signal peptide:

amino acids 1-24

N-glycosylation sites.

amino acids 190-194 and 251-255

Glycosaminoglycan attachment sites.

amino acids 149-153 and 155-159

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 26-30

Casein kinase II phosphorylation sites.

amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343
and 349-353

Tyrosine kinase phosphorylation site.

amino acids 303-310

N-myristoylation sites.

amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and
313-319

Aspartic acid and asparagine hydroxylation site.

amino acids 308-320

EGF-like domain cysteine pattern signature.

amino acids 166-178

Leucine zipper pattern.

amino acids 94-116

11-12-13-14-15-16-17-18-19-20-21-22-23-24-25-26-27-28-29-30-31-32-33-34-35-36-37-38-39-40-41-42-43-44-45-46-47-48-49-50-51-52-53-54-55-56-57-58-59-60-61-62-63-64-65-66-67-68-69-70-71-72-73-74-75-76-77-78-79-80-81-82-83-84-85-86-87-88-89-90-91-92-93-94-95-96-97-98-99-100-101-102-103-104-105-106-107-108-109-110-111-112-113-114-115-116-117-118-119-120-121-122-123-124-125-126-127-128-129-130-131-132-133-134-135-136-137-138-139-140-141-142-143-144-145-146-147-148-149-150-151-152-153-154-155-156-157-158-159-160-161-162-163-164-165-166-167-168-169-170-171-172-173-174-175-176-177-178-179-180-181-182-183-184-185-186-187-188-189-190-191-192-193-194-195-196-197-198-199-200-201-202-203-204-205-206-207-208-209-210-211-212-213-214-215-216-217-218-219-220-221-222-223-224-225-226-227-228-229-230-231-232-233-234-235-236-237-238-239-240-241-242-243-244-245-246-247-248-249-250-251-252-253-254-255-256-257-258-259-260-261-262-263-264-265-266-267-268-269-270-271-272-273-274-275-276-277-278-279-280-281-282-283-284-285-286-287-288-289-290-291-292-293-294-295-296-297-298-299-300-301-302-303-304-305-306-307-308-309-310-311-312-313-314-315-316-317-318-319-320-321-322-323-324-325-326-327-328-329-330-331-332-333-334-335-336-337-338-339-340-341-342-343-344-345-346-347-348-349-350-351-352-353-354-355-356-357-358-359-360-361-362-363-364-365-366-367-368-369-370-371-372-373-374-375-376-377-378-379-380-381-382-383-384-385-386-387-388-389-390-391-392-393-394-395-396-397-398-399-400-401-402-403-404-405-406-407-408-409-410-411-412-413-414-415-416-417-418-419-420-421-422-423-424-425-426-427-428-429-430-431-432-433-434-435-436-437-438-439-440-441-442-443-444-445-446-447-448-449-450-451-452-453-454-455-456-457-458-459-460-461-462-463-464-465-466-467-468-469-470-471-472-473-474-475-476-477-478-479-480-481-482-483-484-485-486-487-488-489-490-491-492-493-494-495-496-497-498-499-500-501-502-503-504-505-506-507-508-509-510-511-512-513-514-515-516-517-518-519-520-521-522-523-524-525-526-527-528-529-530-531-532-533-534-535-536-537-538-539-540-541-542-543-544-545-546-547-548-549-550-551-552-553-554-555-556-557-558-559-560-561-562-563-564-565-566-567-568-569-570-571-572-573-574-575-576-577-578-579-580-581-582-583-584-585-586-587-588-589-590-591-592-593-594-595-596-597-598-599-600-601-602-603-604-605-606-607-608-609-610-611-612-613-614-615-616-617-618-619-620-621-622-623-624-625-626-627-628-629-630-631-632-633-634-635-636-637-638-639-640-641-642-643-644-645-646-647-648-649-650-651-652-653-654-655-656-657-658-659-660-661-662-663-664-665-666-667-668-669-670-671-672-673-674-675-676-677-678-679-680-681-682-683-684-685-686-687-688-689-690-691-692-693-694-695-696-697-698-699-700-701-702-703-704-705-706-707-708-709-710-711-712-713-714-715-716-717-718-719-720-721-722-723-724-725-726-727-728-729-730-731-732-733-734-735-736-737-738-739-740-741-742-743-744-745-746-747-748-749-750-751-752-753-754-755-756-757-758-759-760-761-762-763-764-765-766-767-768-769-770-771-772-773-774-775-776-777-778-779-780-781-782-783-784-785-786-787-788-789-790-791-792-793-794-795-796-797-798-799-800-801-802-803-804-805-806-807-808-809-810-811-812-813-814-815-816-817-818-819-820-821-822-823-824-825-826-827-828-829-830-831-832-833-834-835-836-837-838-839-840-841-842-843-844-845-846-847-848-849-850-851-852-853-854-855-856-857-858-859-860-861-862-863-864-865-866-867-868-869-870-871-872-873-874-875-876-877-878-879-880-881-882-883-884-885-886-887-888-889-890-891-892-893-894-895-896-897-898-899-900-901-902-903-904-905-906-907-908-909-910-911-912-913-914-915-916-917-918-919-920-921-922-923-924-925-926-927-928-929-930-931-932-933-934-935-936-937-938-939-940-941-942-943-944-945-946-947-948-949-950-951-952-953-954-955-956-957-958-959-960-961-962-963-964-965-966-967-968-969-970-971-972-973-974-975-976-977-978-979-980-981-982-983-984-985-986-987-988-989-990-991-992-993-994-995-996-997-998-999-1000-1001-1002-1003-1004-1005-1006-1007-1008-1009-1010-1011-1012-1013-1014-1015-1016-1017-1018-1019-1020-1021-1022-1023-1024-1025-1026-1027-1028-1029-1030-1031-1032-1033-1034-1035-1036-1037-1038-1039-1040-1041-1042-1043-1044-10

AAGGTCCTCAACTGCACCTCGGTTCTATCGATTGAATTCCTCCCGGGGATCCTCTAGAGATCCCTC
 GACCTCGACCCACCGCTCCGACGAGCCGGGAGGCGCAGCCGACCGCTCTAAAGCGGGAACA
 CGCCTGGCTGAGGAGCTCGACGCGCAGCAGATGATCGCGCGCCAGGTTGCTGTAGTGGCG
 GCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGCAATGCCCGGAGGAGCGCCTTC
 CCTGCCCGCGCTCTGGCTCTTGGAGCATCTCTCTGTGCTGCTGCGCAGCTCGGGCGGAGGC
 CGGGCCCGCGCAGGAGGAGCATCTGACTTATGATCGATGCTCAGGACCGAAGAGTACTCA
 TAGGATTTGAAGAAGATATCCTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGA
 TTTAGAAAAGCGCAACAGAGAATGCCAGCTATTCTCTGCAATATCAATTCATGAATTTTAC
 CTGGCAAGCTCGAGGCGAGGACATACTTTATGAATTCCTGTCTGCTGCGCTCCCTGGATA
 AAGGCATCATGCGAGATCAACCCGTCAATGTCCCTCTGCTGGAACAGTGCGCTCAACAGGCA
 TCAGTTGTTCAAGTTGGTTTCCCATGCTTTGGAACACAGGATGGGTGGCAGCATTTGAAGT
 GGATGTGATTGTTATGAATTCGAAAGCAACACATCTTCAAACACCTCAAAATGCTATCT
 TCTTTAAACATGTCAACAGCTGAGTGCGCCAGCGGGTGCCGAATGGAGGCTTTTGTAAAT
 GAAGAAGCATCTGCGAGGTGCTGTATGGGTTTCCAGGCATCACTGTGGAAGAAGCCCTTTG
 TACCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCACCTG
 GATTTATGAGGTGAACCTGTGACAAAGCAAACTGCTCAACACCTGCTTTAATGGAGGACCT
 GTTTTCTACCCCTGGAAGATGATTTGCCCTCAGGACATGAGGGAGAGCAGTGTGAATCAG
 CAAATGCCCAACCCCTGTGCAAAATGGAGGTAAATGTCATTGTGTAAGAAGCAATGTAAGTGT
 TCAAAGGTTACCGGGAGACCTCTGTTCAAAGCCTGCTCGCGAGCTGGTGTGGTGACAT
 GGAACCTGCCATGAACCCAAACAAATGCCAATGTCAAGAAGTTGGATGGAAGACACTGCAA
 TAAAGAGCTACGAAGCGAGCTCATACATGCCCTGAGGCCAGCAGGCCGCCACCTGAGCGAC
 ACACGCTTCACTTAAAGGCCGAGGAGCGGCGGATCCACCTGAATCCAATTACATCTCTGG
 TGAACCTCCGACATCTGAAACGTTTAAAGTTACACCAAGTTCATAGCCTTTGTTAACTTTCA
 TGTGTTGAATGTTCAAATAATGTTCATTACACTACATTAAGAATCATGGCTGAATTTATTAGCT
 TCAATTATAAATCTCAGTGCATATATTACTTCTCTTTTAAAGTTTCTAAGTAGCTGTGAT
 CATGATGGTATAGATTTTCTGTTTCACTGCTTTGGGACAGATTTTATATTATGTCAATTGA
 TCAGTTTAAATTTTCACTGTGTAGTTGGCAGATATTTCAAATATCAATGCAATTTATGGT
 TCTGGGGCGAGGGGAACATCAGAAAGTTAAATTGGGCAGAAAATGCGTAGTCAACGAAGAT
 TTGATGGTGCAGTTAATGTTGAAGTTACAGCAATTTTCAGATTTTATTGTGCATATTTAGAT
 GTTGTGTACATTTTAAAAATGCTCTTAATTTTAAACTCTCAATACAAATATTTTGGAC
 TTACCATTTATTCAGAGATTCAGTATTAATAAAAAAATAAATTAACCTGGTGTGGATTTT
 AAAACAATATAATATTTCTAAACACCAATGAAATAGGAATATAATGTATGAACATTTTTCAT
 TGGCTTGAAGCAATATAATATATTGTAAACAAACACAGCTTTACCTATAAACATTTTAT
 ACTGTTTGTATGTATAAAATAAAGGTGCTGCTTTAGTTTGTGAAAAAATAAATAAATAA
 AAAAAAATAAATAAATAAAGGGCGCGCGCACTTAGAGTCGACCTGCAGAGCTTGGC
 CGCCTAGCGCCCAACTTGTTTATTGCGCTTATAATG

FIGURE 4

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094
><subunit 1 of 1, 379 aa, 0 stop
><MW: 41528, pI: 7.97, NX(S/T): 2
MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGG
MAPFTTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFLLSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNITLQTPQNAIFFKTCQQAECPGGC
RNGGFCNERRICECPDGFGHGPHEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST
TCFNGGTCFYPGKICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAEERRDP
PESNYIW

Signal peptide:

amino acids 1-28

N-glycosylation site.

amino acids 88-92, 245-249

Casein kinase II phosphorylation site.

amino acids 319-323

Tyrosine kinase phosphorylation site.

amino acids 370-378

N-myristoylation sites.

amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).

amino acids 285-293

EGF-like domain cysteine pattern signature.

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

FIGURE 5

CGGACGCGTGGGCGTCCGGCGGTTCGCAGAGCCAGGAGGCGGAGGCGCGCGGGCCAGCCTGGG
CCCCAGCCCACACCTTACCAGGGGCCAGGAGCCACCATGTGGCGATGTCCACTGGGGCTAC
TGCTGTTGCTGCCCGTGGCTGGCCACTTGGCTCTGGGTGCCCAGCAGGGTCTGGGCGCCGG
GAGCTAGCACCGGGTCTGCACCTGCGGGGCATCCGGGACCGGGAGGCCGGTACTGCCAGGA
GCAGGACCTGTGCTGCCGCGGCGGTGCCGACGACTGTGCCCTGCCCTACTGGGCGCCATCT
GTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCTGCCCTGACTTCTGGGACTTC
TGCTCGGGCTGGCAACCCCTTTTCCCCCGATCCAAGGATGTATGCATGGAGGTCGTATCTA
TCCAGTCTTGGGAACGCTACTGGGACAACTGTAACCGTTGCACCTGCCAGGAGAACAGGCAGT
GGCATGGTGGATCCAGACATGATCAAAGCCATCAACCAGGGCAACTATGGCTGGCAGGCTGG
GAACCACAGCGCCTTCTGGGGCATGACCCTGGATGAGGGCATTTCGCTACCGCCTGGGCACCA
TCCGCCCATCTTCTCGGTTCATGAACATGCATGAAATTTATACAGTGCTGAACCCAGGGGAG
GTGCTTCCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCCAACCTGATTCTATGAGCCTCTTGA
CCAAGGCAACTGTGCAGGCTCCTGGGCCCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCT
CAATCCATTCTCTGGGACACATGACGCCTGTCTGTGCCCCAGAACCTGCTGTCTTGTGAC
ACCCACCAGCAGCAGGGGCTGCCGCGGTGGGCGTCTCGATGGTGCCCTGGTGGTTCTTGCCTCG
CCGAGGGGTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACAGAGACGAGGCTGGCC
CTGCGCCCCCTGTATGATGCACAGCCGAGCCATGGGTGGGGCAAGCGCCAGGCCACTGCC
CACTGCCCCAACAGCTATGTTAAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCT
CGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCA
TGGAGGTGCATGAGGACTTCTTCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGC
CTTGGGAGGCCAGAGAGATACCGCGGCATGGGACCCACTCAGTCAAGATCACAGGATGGGG
AGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGCGGCCAACTCTGGGGCC
CAGCCTGGGGCGAGAGGGGCCACTTCCGCATCGTGCGCGGCGTCAATGAGTGCGACATCGAG
AGCTTCTGTCTGGGCGTCTGGGGCCGCGTGGGCATGGAGGACATGGGTCACTACTGAGGCTG
CGGGCACCACGCGGGGTCCGGCTCTGGGATCCAGGCTAAGGGCCGGCGGAAGAGGCCCAATG
GGGCGGTGACCCAGCCTCGCCCGACAGAGCCCGGGGCGCAGGCGGGCGCCAGGGCGCTAAT
CCCCGCGCGGTTCCCGTGACGCAGCGCCCCGCTGGGAGCCGCGGGCAGGCAGACTGGCG
GAGCCCCAGACCTCCAGTGGGGACGGGGCAGGGCCTGGCTGGGAAGAGCAGACGTGACG
ATCCAGGCCTCTGGCGCCCCCACTCAAGACTACCAAAGCCAGGACCTCAAGTCTCCAG
CCCAATACCCACCCCAATCCCGTATCTTTTTTTTTTTTTTTTATAGCAGGGTCTTGCTCCG
TTGCCAGGTTGGAGTGACGTGGCCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTTCA
AGTGACCCCTCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCACCACCACACTGGC
TAATTTTTGTATTTTTGTAAAGAGGGGGGTCTCACTGTGTGCCAGGCTGGTTTCGAACT
CCTGGGCTCAAGCGGTCAACCTGCCTCCGCCCTCCAAAGTGCTGGGATTGCAGGCATGAGCC
ACTGCACCAGCCCTGTATTCTTATTCTTCAGATATTTATTTTCTTTCACTGTTTTTAAAA
TAAACCAAAGTATTGATAAAAAAAA

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

FIGURE 6

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223
><subunit 1 of 1, 164 aa, 1 stop
><MW: 18359, pI: 7.45, NX(S/T): 1
MWRCPLGLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQLCCRGRADDC
ALPYLGAICYDLFCNRTVSDCCPDFWDFCLGVPPFPPIQGCMHGGRIYVVLGTYWDCNR
CTCQENRQWHGGSRRHDQSHQPGQLWLAGWEPQRLLGHDPG
```

N-glycosylation site.

amino acids 78-82, 161-165

Casein kinase II phosphorylation site.

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300,
411-415

N-myristoylation site.

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230,
269-275, 378-384, 442-448

Amidation site.

amino acids 26-30, 318-322

Eukaryotic thiol (cysteine) proteases histidine active site.

amino acids 398-409

FIGURE 7

AGGCTCCTTGGCCCTTTTCCACAGCAAGCTTNTGCNATCCCGATTGTTGTCTCAAATCCA
ATTCTCTTGGGACACATNACGCCTGTCCTTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC
CAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCTTGCGTCGCCGAGG
GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC
CCCCCTGTATGATGCACAGCCGAGCCATGGGTGCGGGCAAGCGCCAGGCCACTGCCCACTGC
CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTC
CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGG
TGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGG
AGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAG

100
90
80
70
60
50
40
30
20
10
0
A
C
G
T
-
100
90
80
70
60
50
40
30
20
10
0
A
C
G
T
-
100
90
80
70
60
50
40
30
20
10
0
A
C
G
T
-

FIGURE 8

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGTGCTACT
CCTGCAAAGCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG
GAGCAGTGCTGGACCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG
CAGCTTGAATGCGTGGATGACTCACAGGACTACTACGTGGGCAAGAAGAACATCACGTGCT
GTGACACCGACTTGTGCAACGCCAGCGGGGCCATGCCCTGCAGCCGGCTGCCGCCATCCTT
GCGTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGACCCGGCCAGCTATAGGCTCTGGGGGG
CCCCGCTGCAGCCCACACTGGGTGTGGTGCCCCAGGCCCTCTGTGCCACTCCTCACAGACCTG
GCCAGTGGGAGCCTGTCCTGGTTCCCTGAGGCACATCCTAACGCAAGTCTGACCATGTATGT
CTGCACCCCTGTCCCCCACCCTGACCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA
TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCTCCAACCCTCTCTGCTGCTGTTTC
CATGGCCAGCATTCTCCACCCTTAACCCTGTGCTCAGGCACCTCTTCCCCCAGGAAGCCTT
CCCTGCCCCACCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCTCCCGCAGCCAGCA
GGGGACAGGCACTCAGGAGGGGCCAGTAAAGGCTGAGATGAAGTGGACTGAGTAGAACTGGA
GGACAAGAGTCGACGTGAGTTCTTGGGAGTCTCCAGAGATGGGGCTGGAGGCCTGGAGGAA
GGGGCCAGGCCTCACATTCTGTGGGGCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT
AATAAACACCTGTTGGATAAGCCAAAAAA

1994-1995

Signal peptide:

amino acids 1-47

N-glycosylation site.

amino acids 31-35, 74-78, 84-88

Casein kinase II phosphorylation site.

amino acids 22-26, 76-80

N-myristoylation site.

amino acids 56-60

Amidation site.

amino acids 70-74

FIGURE 10

CCCACGCGTCCGAACCTCTCCAGCGATGGGAGCCGCCCGCCTGCTGCCAACCTCACTCTGT
GCTTACAGCTGCTGATTCTCTGCTGTCAAACCTCAGTACGTGAGGGACCAGGGCGCCATGACC
GACCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCA
CGTGCAAGTCAACGGGCGTTCGCATCTTCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAAGC
TCATAGTGGAGACGGACAGTTTGGCAGCCGGTTTCGCATCAAAGGGCTGAGAGTGAGAAG
TACATCTGTATGAACAAGAGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTG
CGTGTTCACGGAGATCGTCTGGAGAACAACTATACGGCCTTCAGAACGCCCGGCACAGAGG
GCTGGTTCATGGCCTTCACGCGGCAGGGGCGGCCCGCCAGGCTTCCCGCAGCCGCCAGAAC
CAGCGCGAGGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCCAGC
CGAGAAGCAGAAGCAGTTTCGAGTTTGTGGGCTCCGCCCCACCCGCCGACCAAGCGCACAC
GGCGGCCCCAGCCCCCTCACGTAGTCTGGGAGGCAGGGGGCAGCAGCCCCCTGGGCCGCCTCCC
CACCCCTTTCCTTCTTAATCCAAGGACTGGGCTGGGGTGCGGGAGGGGAGCCAGATCCCC
GAGGGAGGACCTGAGGGCCGCGAAGCATCCGAGCCCCAGCTGGGAAGGGGCAGGCCGGTG
CCCCAGGGGCGGCTGGCAGAGTCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAAC
TCTGCTTCTCGGATCTCCCTCAGTCTGCCCCAGCCCCCAAACCTCCTCTGGCTAGACTGTA
GGAAGGACTTTTGTGTTTGTGTTTGTGTTTTCAGGAAAAAAGAAAGGGAGAGAGGAAAAATAG
AGGGTTGTCCACTCTCACATTCACGACCCAGGCCTGCACCCACCCCCAACCTCCAGGCC
CGAATAAAACCATTTTCCTGC

FIGURE 11

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRI
SATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNRGKLI GKPSGKSKDCVFTEIVLE
NNYTAFQNA RHEGWMAFTRQGRPRQASRSRQ NQREAHFIKRLYQGQLPFPNHA EKQKQFEF
VGSAPTRRTKRTRRPQPLT

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

FIGURE 12

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATCTCCCTGTTGAATTTTTTGCACATGGAG
GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC
CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTCTTGGACTTCAACAGAACCCCATCCAGT
CATTTTGATTTTGTCTGTTTATTTTTTTTTCTTTTCTTTTCCACCACATTGTATTTTAT
TTCGTACTTCAGAAATGGGCCTACAGACCACAAAGTGGCCAGCCATGGGGCTTTTTCTCT
GAAGTCTTGGCTTATCATTTCCTGGGGCTCTACTCACAGGTGTCCAACTCCTGGCCCTGCC
CTAGTGTGTGCCCTGTGCACAGGAACCTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG
CCTCTTGGGATCCCGAGGGCGTAACCGTACTCTACCTCCACAAACCAAAATTAATAATGC
TGGATTTCTCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC
AACTGGACGAATTCGCCATGAACCTTCCAAGAATGTACAGAGTTCCTCCATTTGCAGGAAAAAC
AATATTACAGACCATTTACGGGCTGCTCTTGCCAGCTCTTTGAAGCTTGAAGAGCTGCACCT
GGATGACAACCTCCATATCCACAGTGGGGGTGGAAGACGGGGCTTCCGGGAGGCTATTAGCC
TCAAATGTTGTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCTCTGTGGAC
TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTATATCCGACATGGCCTTCCAGAA
TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG
AGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTGCC
CACCTCTCTCCGATCTCCAGGTACGCATCTGATCAGGCTCTATTGTGCAGGACAACCAGAT
AAACCACATTCCTTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACCGCTGGATATATCCA
ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC
ACTGCTCGGAATAACCCCTGGTTTTGTGACTGCAGTATTAAATGGGTACAGAAATGGCTCAA
ATATATCCCTTCTCATCTCTCAACGTGCGGGGTTTTCATGTGCCAAGGTCTCGAACAGTCCGGG
GGATGGCCGTGAGGAAATTAATATGAATCTTTTGTCTGTGCCACACAGACCCCGGCCCTG
CCTCTCTTACCCACAGCCCCAAGTACAGCTTCTCCGACCATCAGCCTCCACCCCTCTCTAT
TCCAAACCTAGCAGAAGCTACACGCCCTCAAACCTCTACCATCGAACTTCCACGATTC
CTGACTGGGATGCGAGAGAAAGTGACCCACCTATTCTGAACGGATCCAGCTCTCTATC
CATTTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTACCGGTGATGGCATA
CAAACCTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTACAGGAGCGCATAG
TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCGATCCACCTATCGGATT
TGTTTAGTGCACCTGGATGCTTTTAACTACCGCGCGGTAGAAGACCAATTTGTTACAGAGC
CACCACCCATGCCTCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA
CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATT
GTGCTGGTGGTCTTGTCTAGCGTCTTTTGTGTCATATGCACAAAAGGGGCGCTACACCTC
CCAGAACTGGAAATACAACCGGGGCGGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGA
AGGCAACTCTCATCTCTGGAGATGACAGAAACAGTTTTTCAGATCGTCTCTTAAATAACGAT
CAACTCTTAAAGGAGATTTACAGCTGACGCCATTACACCCAAATGGGGGCATTAAATTA
CACAGACTGCCATATCCCCAAACAACATGCGATACTGCAACAGCAGCGTGCAGACCTTGAGGC
ACTGCCATACGTTAGCCCGAGAGGCCAGCGTTATCAAGGCGCAATTTAGACTCTTGAGAA
CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT
TTGTGCATTGTAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAGTGT
CTATCTTTTCTATTCAAGTTAATACAAACAGTTTTGTAACTCTTTGCTTTTAAATCTT

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

FIGURE 13

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDNRFVYCNERSLTSVPLGIP
EGVTVLXLHNNQINNAGFPAELHNQSVHTVLYLGNQLDEFFPMNLPKNVRLHLQENNIQTI
SRAALAQLLKLLEELHDDNSISTVGVEDGAFREAI SLKLLFLSKNHLSSVPVGLPVDLQELR
VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD
LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN
PWFCDCSIKWTEWLKYPSSLNVRGFMCGPEQVRGMAVRELNMMNLLSCPTTTPGLPLFTP
APSTASPTTQPTTSLIPNPSRSYTPPTPTTSKLPITPDWDGREVRTPPISERIQLSIHFVND
TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL
DAFNRYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSMSMGSPFLLAGLIGGAVIFVLVVL
LSVFCWHMHKKGRYTSQKWYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG
DFRLQPIYTPNGGINYTDCHI PNNMRYCNSSVPDL EHCHT

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,
522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

FIGURE 14

ACTTGGGACAAGCGCGCGCGGAGACAGAGGCGAGAGCAGAGCTGGGGCTCCGTCTCGCCTCCACGAGCG
ATCCCCGAGGAGAGCCGCGGCCCTCGGCGGAGGCGAAGAGGCCGACGAGGAAGACCCGGGTGGCTGCGCCCTGCC
TCGCTTCCAGGCGCGCGCGGCTGACGCTTGCCTCTTCTCGCTCGCCTTGAAAATGGAAAGAATGCTCGCAGGCT
GCTTTCTGTGATCTCTGGACAGATCGTCTCTCTCTCGCGAGGCGAGGAGCGGTACGCTGGGAGGTCCATCT
CTAGGGGCGACACGCTCGGACCCACCCGCGAGCGGCCCTCTCGGAGATTTCTGTGAGAACCAAGCGGCGAGACC
TGGTTTTCATCATTTGACAGTCTTCGCACTGCAACACCCATGACTAGTCAAAAGGTTCAGGAGTTTCATCGTGGAC
TCTTGCAATTTCTTGACATTTGGTCTGTATGTCACCCGAGTGGGCTGTCCAAATGGGAGGTTTCATCGTGGAC
AGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCGAGGTGGAGCGTGTGTCAAGAGGATGCGGCACTGTGTCCA
GCACCATGACTTGGCTGGCCATCCAGTATGCCCTGAAACATCGCATTCAGAAAGCAGAGGGGGCGCGCCCTGA
GGGGAATGTGCGCACGGGTACATAATGATCTGTGACAGATGGGAGACCTCAGGAGCTCCGTGCGGAGGTGGCTGCTA
AGGCAACGGGACACGGGCATCTTAATCTTGGCCATTGGTGTGGGCGAGGTAGACTTCAACACCTTGAAGTCCAATTG
GGAGTGAAGCCCATGAGGACCATGTCTTCTTGTGGCCAAATTCAGCGAGATTGAGACGCTGACCTCCGTGTTC
AGAAGAAGTTGTGCAACGGGCCACATGTGCGAGCACCCCTGGAGCATAACTGTGCCCACTTTCGATCAACATCCCTG
GCTCATACGTCTGCGAGGTGCAACAGGCTACATTTCTCAACTCGGATCAGACGACTTGCAGAAATCCAGGATCTGT
GTGCGCATGGAGGACCAACTGTGAGCAGCTCTGTGTGAAATGTGCGGGCTCTTTCGTCTGCGAGTGTACAGT
GTACGCGCTTGGCTGAGGATGGGAAGGAGGTGTGTGCTGTGGACTACTGTGCTCAGAAAACAGGATGTGAAC
ATGAGTGTGTAATGTGTATGCTCTTACCTTTGCGAGTGAAGGATTTGCTCTTAACCCAGATGAAAAA
CGTGCACAAGGATCAACTACTGTGCACTGAACAAACCGGGCTGTGAGCATGAGTGCCTCAACATGGAGGAGAGCT
ACTACTGCCGTGCCACCGTGGCTACACTCTGGACCCCAATGGCAAAACCTGCGAGCGAGTGGACCACTGTGCAC
AGCAGGACCATGGCTGTGAGCAGCTGTGTGTGAACACGGAGGATTCCTTCGTCTGCCAGTGTGCAAGGCTTCC
TCATCAACGAGGACCTCAAGACCTGCTCCCGGGTGGATTACTGCTGTGAGTGACCATGGTTGTGTAATCTCTGT
GTGTCAACATGGACAGATCTTTGCTGTGAGTGTCTGAGGACACGCTGCTCGCAGCGATGGGAAGACGCTGTGT
CAAAATTTGGAATCTTTGTGCTCTGGGGGACCAAGGTTGTGAACATTCGTGTGTAAAGCAGTGAAGATTCGTTGTGT
GCCAGTGTCTTTGAAGGTTATATACTCCGTGAAGATGGAAAAACCTGCAAGGAAGAAGATGTCTGCGAAGCTATAG
ACCATGGCTGTGAACACATTTGTGTGAACAGTGAACGACTCATACAGTGGCAGTGTCTGGAGGATTCGGCTTCG
CTGAGGATGGGAACGCTGCCGAAGGAAGGATGTGTGCAAAATCAACCCACCATGGCTGGGAACACCTTTGTGTTA
ATAATGGGAATTCCTACATCTGCAATGTCTAGAGGATTTGTTCTAGCTGAGGACGGAAGACGCTGCAAGAAT
GCACTGAAGGCCAATTTGACCTGGTCTTTGTGATCGATGGATCCAAAGATCTGGAGAGAGAAATTTGAGGTCTG
TGAAGCAGTTTGTCTGGAATTTAGATTTCTTGACAAATTTCCCAAGCCGCTCAGAGTGGGCTGCTTCAAGT
ATTCACACAGGTCCACACAGATTTCACTCTGAGAACTTCACTCAGGCCAAGACGATGAAAAAGCGGTGGCC
ACATGAAATACATGGGAAAGGCTCTATGACTGGCTGGGCCCTGAAACACATGTTTTCAGAGAAGTTTTCACCAAG
GAGAAGGGGCCAGGCCCTTTCCACAGGGGTGCCAGAGCGGCATTGTGTTCACCGACGAGCGGCTCAGGATG
ACGTGCTGGTGGGCGCTAAAGCAAGGCCAATGGTATCACTATGTATCTGTCTGGGTAGGAAAGACGATG
AGGAGGAATACAGAGAGATTGCTCTGAGCCCAACAAAGCATCTCTCTATGCGCAAGACTTCACGCAATGG
ATGAGTAACTGTGAAAACTCAAGAAAGGCATCTGTGAAGCTCAGAGAAGCTCAGGATGGAGGACCTCCAG
CAGGGGAATGCGCAAAACCGGTCCACAGCCAAACAGAAATCTGAGCCAGTACCATAAATATCCAAAGCTTACTT
CTTCTTCAATTTTTCAGTGTCTGACGATCTACAGATATCTGTTTGAAGAAGACAATCTTTACGGTCTACACAAAGCTT
CCCATTCACAAACCTTTCAGGAAGCCCTTTGGAAGAAAAACGATCAATGCAAAATGGAAACCTTATAATGT
TCCAGAACCTTCCAAACCAAGAAATGAAGAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCC
TGGAATATCGCTGAGATACAGATGAAGATTAGAATCGCGACACATTTGTAGTCACTGTATACCGGATTAACA
GAACGAGTGCAGAGCCCAAGCTCTCAGGCTATTGTAATCAATAATGTTGTGAAGTAAAAACATCAGTACTGA
GAACCTGTGTTTSCCAACAGCAAGACAAAGAGTATACACTAACTGTATAAATTTATCTAGGAAAAAATCTCT
TCAGAAATCTTAGATGAATTTACAGGTGAGAATGAATAGCTATGCAAGGATTTTGTATATATACTGTGGACAC
AATCTGCTTCTGCTCATCTCGCTTGTGTGCAATCTCATTTGACTATACGATAAGATTTGCAAGCTTTACTT
CTGTAGAACATCGGCATAGGAAATGCTGTTTTTTGTACTGGACTTACCTTGATATATGATATGGATGATG
CATAAAAATCAGGACATATGTACTGTGGAAACAGTTGGATTTTTTATACAATATTAATAATCCACACTTCAG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 15

MEKMLAGCFLLLILGQIVLLPAEARERSRGRSISRGRHARTHPTQALLESSCENKRADLVFII
DSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV
KRMRLHSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARD
TGILIFAIQVGVQVDFNTLKSIGSEPHEDHVFVANFSQIETLTLSVFQKKLCTAHMCSTLEHN
CAHFCINIPGSYVCRCKQGYILNSDQTTTCRIQDLCAMEDHNCBQLCVNVPGSFVCQCYSGYA
LAEDGKRCVAVDYCASNHGCEHECVNADGSYLCQCHEGFALNPDEKTCCTRINICALNKPGE
EHECVNMEESYCRCHRGYTLDPNGKTC SRVDHCAQQDHGCEQLCLNTEDSFVCQCSGFLI
NEDLKTCSRVDYCLLSDHGCEYSCVNMDRSFACQCPGEGHVLRS DGKTC AKLDS CALGDHGCE
HSCVSSSEDSFVCQCFEGYILREDGKTCRRKDVCAIDHGCEHICVNSDDSYTCECLEGFRLA
EDGKRRCRKDVCKSTHHGCEHICVNNNGNSYICKSEGFVLAEDGRRCKKCTEGPIDLVFVID
GSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNNSAKDMKKAVA
HMKYMGKSGMTGLALKHMFERSFTQGEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN
GITMYAVGVGKAIEEELQEIASEPTNKHLYFAEDFSTMDIEISEKLLKGICAELESDSDGRQDS
PAGELPKTVQQTPESEPVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPLE
EEKHDQCKCENLIMFONLANEEVRKLTQRLEEMTQRMALENRLRYR

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 115-119, 606-610, 892-896

Casein kinase II phosphorylation site.

amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247,
401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784,
781-785, 819-823, 866-870

N-myristoylation site.

amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500,
639-645, 690-696, 752-758, 792-798

Amidation site.

amino acids 314-318, 560-564, 601-605

Aspartic acid and asparagine hydroxylation site.

amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464,
540-546, 581-587

FIGURE 16

GGAGCCGCCCTGGGTGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTCGCGCAGCCTCG
GCACCTGTCAGGTCCGTGCGTCCCGCGGCTGGCGCCCTGACTCCGTCCCGGCAGGGAGGGC
CATGATTTCCTCCCGGGGCCCCCTGGTGACCAACTGTGTCGGTTTTTGTTCCTGGGGCTGA
GTGCCCTCGCGCCCCCTCGCGGGCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG
GCGGTGGAGGGAGGGGAAGTGGTGCTTCACAGCGTGTACACCTTGCACGGGGAGGTGTCTTC
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAACAGAAAGAAAAGGAGGATC
AGGTGTTGTCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT
TAGAACTCAATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT
GTGGGGGCAAAACGTGACCTTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCCTTGCAACCAGCATTAGATGTATCCGTG
GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCAC
AATGAGGTGGGCACTGCCAATGTAATGTGACGTGGAAGTGAGCACAGGGCCTGGAGCTGC
AGTGGTTGTCTGGAGCTGTTGTGGGTACCCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC
TCTTGTACCACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC
ATTGCTCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT
TTCCTCTGTACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCAGGCCTGGTGCAT
TGACCCCAAGCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCACGACAGAT
GGGGCCACCCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTCTTCCTCTGGCTTGAGCCG
CATGGGTGCTGTGCCTGTGATGGTGCCTGCCAGAGTCAAGCTGGCTCTCTGGTATGATGAC
CCCACCACTATTGGCTAAAGGATTTGGGGTCTCTCCTTCTTATAAGGGTCACCTCTAGCAC
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCCACCTCTC
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG
CTGAAATTAGCTACTCACCAAGAGTGAGGGGCAGAGACTTCAGTCACTGAGTCTCCAGGC
CCCCCTGATCTGTACCCACCCCTATCTAACACCACCTTGGCTCCCACTCCAGCTCCCTGT
ATTGATATAACCTGTCAAGCTGGCTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC
TTATTAATAAATAACATGAAATATGTGTTGTTTTTCATTGCAAAATTTAAATAAAGATACATAA
TGTTTGTATGAAAAA

FIGURE 17

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS
SQPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKGSGPY
SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPVHGAVNLTSCQSPRSKPAVQYQ
WDRQLPSFQTFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLLEVSTGPGAA
VVAGAVVGTTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLWPWKSSDTISKNGTL
SSVTSARALRPPHGPPRGALTPTPSLSSQALPSPRLPTTDGAHPQPISSIPGGVSSSGLSR
MGAVPVMVPAQSQAQGLV

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,
262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 164-175

THE UNIVERSITY OF CHICAGO

CGCCACCACCTGCGGCCACCGCCAATGAACGCGCTCCCGCTCCTAGTGGTTTTTTTCCACTTTGG
TTGAATTTGTTCTCTATCTCAAAATTGACCACGACAGCTTTGCTCCCAAATGCAGAAATGTGTA
AATACGCAATGGAAATGAAGCGCTGCTATTGCAACATGGGATTTTTCAGGAAATGGTGTACAA
TTTGTGGAAGATGATAATGAATGTGGAATTTAACTCAGTCTGTGGCGGAAATGCTAATTTGC
ATAACACAGGAGGAAGTTATTATGTATGTGTGTACCTGGCTTCAGATCCAGCAGTAAACCA
AGCAGGTTTATCCTAATAAGTGAAGACCGTCTGTATAGAAAATGGTGAATGCAAACTGCCATT
TAGATAATGTCTGTATAGCTGCAAAATATTAAATAAACTTTAACAAAAATCAGATCCATAAAA
GAACCTGTGGCTTTTGCTACAAGAAGCTCTATAGAAAATCTGTGACAGATTTTTCACCAACAG
TATAATTACATATATAGAAATATTAGCTGAATCATCTTTCTATTACAGTTTACCAAGACAACA
CTATCTCAGCCAGGACACCCCTTTCTAACTCAACTTACTGAAATTTGTAAAAACCGTGAAT
AATTTTGTCTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATAGGGAAC
ACATCTTACAAAACCTCATGCACACTGTTGAACAAAGCTCATTTAAGGATATCCGCAAGCTTCC
AAAGAGCCACAGAGTTTATACAAAATTTCAACGGATATAGCTCTCAAAGTTTCTTTTGTGAT
TCATATAACATGAACATATTCTCCTCATATGAATATGGATGGAGACTACATAAATATATT
TCCAAAGAGAAAAGCTGCATATGATTCAAATGGCAATGTTGCAGTTGCATTTTATATATATA
AGAGTTATGGCTCTTTGCTTTTCATCTGACAACTCTTAATTGAAACCTCAAAATATATGAT
AATTCGAAAGGAGGAAAAGAGTCATCTTCAGTAATTTAGTCTCAATGAGCTCAAAACCC
ACCCACATTTATGAATCTGAAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATA
GGTATAGGATATATGTGTCATTTTGGAAATTTACTCACTGATACCATTAAGGAGGCGAGCTGGTCT
TCAGAGGGCTGTGAGCTGACATCTCAAAATGAGACCCACCTCGAGCCGCTGTAATCACTT
GACACATTTTGCAATTTGATGTCTCTGGTCTCTTCATTTGGTATTTAAGCATTTAATATATTC
TTACAAGGATCTCACTCAACTAGGAATAATATTTCTAGTATTTGTCTTGCCATATGCAATTTT
ACCTCTCGTGTCTTTCAGTGAATTTCAAAGCACCAGGACAACAATTCACAAAAATCTTTGCTG
TAGCCATTCTTGTCTGAACTTGTCTTTCTTGTGGGATCAATAACAAATCAATAAGCTCT
TCTGTTCAATCATTTGCCCGAGTGCTACACTACTCTTTTAGTCTGTCTTTCAGTGGATGTC
ATTGAAGGCATACATCTCTATCTCATTTGTGTGGGTGCTCATGTAACAAGGGATTTTGTGA
CAAGAATTTTATATCTTTGGCTATCTAAGCCAGCGCTGTGTAGTTGGATTTTCGGCAGCAC
TAGGATACAGATATTATGGCAACAACCAAGTATGTGGCTTAGCACCGAAGAAACACTTTATT
TGGAGTTTTATAGGACAGCAGTGCCTCATCATCTCTGTGAATCTCTGGCTTTTGGAGTCAT
CATATACAAAGTTTTCGTCACTCAGGGGTGAACCAAGAGTTAGTTGCTTTGAGAACA
TAAGGTCTTGTGCAAGAGGAGCCCTCGTCTCTCGTTCTCTCGGCACCACCTGGATCTTT
GGGTTCTTCATGTGTGTGCAGCATCAGTGGTTACAGCTTACCTCTTACAGTCAGCAATGC
TTTCCAGGGGATGTTCAITTTTTTATCTCTGTGTGTTTCTTAGAAGATCAAGAAGAAAT
ATTACAGATTTGCTCAAAAATGTGCCCTGTGTTTGGATGTTTAAGGTAACATAGAGAATG
GTGGATAATTTCAACTGTCACAAAAATAAAATTTCAAGCTGTGGATGACCAATGATATAAAAA
TGACTCATCAAAATTTATCCAATTATTAACTACTAGCAAAAAAGTATTTAAATCAATTTTCT
GTTTATGCTATAGGAACGTGTAGATAATAAGGTAAAAATTTATGATCATATAGATATACATGT
TTTTCTATGTGAATAGTTCTGTCAAAAATAGTATGCGACATATTGGAAGATGATTTGGTTT
CTCAGGAGTGATATCACTGCACCAAGGAAGAGATTTTCTTTCAACACGAGAAGTATATGAA
TGTCTGTAAGGAAGAACCTCGCTTGATATTTCTGTAGCTCGTGTGCCTTTGAACATAGTCC
CTACACCCTCGTAATAGAGCTCTCATACAGAAGAGTGAACATAGAAGATGAAGGGGCAGA
ATATCAAAACAGTGAAGGGAAGTATAGATGTTATTTGAATGAATCTGTTTTTCTGTAGAC
TAGCTGAGAAATTTGTGACATAAAATAAAGAATTGAAGAAACACTTTTACATTTTGTGAA
TTGTTCTGAACTTAAATGTCCACTAAAAACACTTAGACTCTGTTTGTCTAAATCTGTTCTT
TTTTCTAATTTCTTAAAAAATAAAAAAGGTTTACCTCCACAAATGAAAAAATAAAAAA
AAAAAATAAAAAAATAAAAAAATAAAAAA

FIGURE 19

MKRLPLLVVVSTLLNCSYTNQCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTTICEDDNEC
GNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVC IENVNANCHLDNVCIAA
NINKTLTKIRSIKEPVALLQEVYRNSVTDLSPTDIIITYIEILAESSLLGYKNNTISAKDTL
NSTLTLEFVKTVNNFVQDRTFVVDKLSVNHRRTHLTKLMHTVEQATLRISQSFOKTTEPDT
NSTDIALKVFFFDSYNMKHIHPHMNDGDYINIFPKRKAAYDSNGNVAVAFLYYKSIGPLLS
SSDNFLLKPKQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHRKVTDYRSLCAF
WNYSPDTMNGSWSSEGCETYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLG
IIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFCSSIIAGL
LHYFFLAFAFWMCIEGIIHLYLIVVGVYKNGFLHKNFYIFGYLSPAVVVGFSAAALGYRYGT
TKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVRHTAGLKPEVSCFENIRSCARGA
LALLFLLGTTWIFGVLHVHASVVTAYLFTVSNFAQGMFIFLFLCVLSRKIQEEYYRLFKNV
PCCFGCLR

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636,
648-664

N-glycosylation site.

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181,
188-192, 249-253, 381-385, 395-399

Glycosaminoglycan attachment site.

amino acids 49-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 360-364

Casein kinase II phosphorylation site.

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154,
155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329,
346-350, 608-612

Tyrosine kinase phosphorylation site.

amino acids 36-44, 669-677, 670-678

N-myristoylation site.

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394,
434-440, 480-486, 521-527

Aspartic acid and asparagine hydroxylation site.

amino acids 75-87

FIGURE 20

TGGAACATATCCTCCCTCATATGAATATGGATGGAGACTACATAAAATATATTTCCAAAGNG
AAAAGCCGGCATATGGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATTATAAGAGTAT
TGGTCCCTTTGCTTTTCATCATCTGACAACTTCTTATTGAAACCTCAAAATTATGATAATTCT
GAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTAGTCTCAATGAGCTCAAACCCACCCAC
ATTATATGAACTTGAAAAATAACATTTACATTAAAGTCATCGAAAGGTCACAGATAGGTATA
GGAGTCTATGTGGCATTTTGGAATACTCACCTGATACCATGAATGGCAGCTGGTCTTCAGAG
GGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACA
TTTTGCAATTTTGATGTCTCTGGTCCTTCATTGGTATTAAAGATTATAATATTCTTACAA
GGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTACCTTC
TGGTTCCTCAGTGAAATTCAAAGCACCAGGA

FIGURE 21

GCTCCAGCCAAAGAACCTCGGGGCCGCTGCGCGGTGGGGAGGAGTTCCCGAAACCCGGCCG
 CTAAGCGAGGCCTCTCTCCCGCAGATCCGAAACGGCTGGGCGGGGTCAACCCGGCTGGGA
 CAAGAAGCCGCCCTGCTGCTGCCCGGCCCGGGGAGGGGCTGGGGCTGGGGCCGAGGCGG
 GGTGTGAGTGGGTGTGTGCGGGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG
 TGTCTTGGGCACCTACCCGTGGGGCCCGTAAGGCGCTACTATATAAGGCTGCCGCCCCGGAG
 CCGCCGCGCCTTCAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC
 GGCACTCAGACCCCGCAGCGCATCCCGGTGCGCGCCAGCTCCCGCACCCCCATCGCCGG
 AGCTGCGCCGAGAGCCCCAGGGAGGTGCCATGCGGAGCGGGTGTGTGGTGGTCCACGTATGG
 ATCTTGGCCGCTCTGGCTGGCCGTGGCCGGGGCGCCCTCGCCTTCTCGGACGCGGGGCC
 CCACGTGCACTACGGCTGGGGCGACCCCATCCGCTGCGGCACCTGTACACCTCCGGCCCCC
 ACGGCTCTCCAGCTGCTTCCCTGCGCATCCGTGCGGACGGCGTGTGGACTGCGCGCGGGGC
 CAGAGCGCGCACAGTTTGTCTGGAGATCAAGGCAGTCTGCTCTGCGGACCGTGGCCATCAAGGG
 CGTGACAGCGTGCCTGCTGTCATGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGT
 ACTCGGAGGAAGACTGTGCTTTTCGAGGAGGAGATCCGCCAGATGGCTACAATGTGTACCGA
 TCCGAGAAGCACCCGCTCCCGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA
 CAGAGGCTTTCTTCACTCTCTCATTTCTGCCATGCTGCCCATGGTCCCAGAGGAGCCTG
 AGGACCTCAGGGGCCACTTGGAAATCGACATGTTCTCTTCCGCCCTGGAGACCGACAGCATG
 GACCCATTGGGCTTGTCAACCGGACTGGAGGCCGTGAGGAGTCCAGCTTTGAGAAGTAACCT
 GAGACCATGCCCGGCCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGACG
 TGCTTCTACAAGAACAGTCTCTGAGTCCACGTTCTGTTTGTAGCTTTAGGAAGAAACATCTAGAA
 GTTGTACATATTCAGAGTTTTCATTTGGCAGTGCCAGTTCTTAGCCAAATAGACTTGTCTGTAT
 CATAACATTGTAAGCCTGTAGCTTGGCCAGCTGCTGCTGGGCCCCCATTTCTGCTCCCTCGA
 GGTGTGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC
 TCACCTTCTTTGGAAAAATCTTATGTCAAGCTGAAATCTCTAATTTTCTCATCACTTC
 CCCAGGAGCAGCCAGAAGCAGGCAGTAGTTTTAATTTTCAGGAACAGGTATCCACTCTGTGA
 AAACAGCAGGTAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG
 ACCATTTTGCCTTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGCATGGCCACCGAG
 GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACTTGAGAATTTCCCC
 CTGAGGCCAGTTCTGTCAATGGATGCTGTCTTGAGAATAAATTTGCTGTCCCGGTGTCACTGCT
 TTCCATCTTCCAGCCACAGCCCTCTGCCACCTCACATGCTTCCCATGGATGGGGCCT
 CCCAGGCCCCCACCCTTATGTCAACCTGCACCTTCTGTTCAAAAATCAGGAAAAAGAAAGAT
 TTGAAGACCCCAAGTCTTGTCAATAACTTGTGTGTGGAAGCAGCGGGGGAAGACCTTAGAAC
 CCTTTCCCGACACTTGGTTTTTCAACATGATATTTATGAGTAATTTATTTTGTATGTACAT
 TCTCTTATTTTCTTACATTATTTATGCCCCAAATATATTTATGTATGTAAGTGAGGTTTG
 TTTGTATATTAATAATGGAGTTTGTGTG

FIGURE 22

MRSGCVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWGDP IRLRHLYTSGPHGLSSCFLRI
RADGVVDCARGQSAHSLSLEIKAVALRTVAIKGVHSVRYLCMGADGKMQGLLQYSEEDCAFEE
EIRPDGYNVYRSEKHLPLVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD
MFSSPLETDSMDPFGLVLTGLEAVRSPSF EK

Signal peptide:

amino acids 1-22

Casein kinase II phosphorylation site.

amino acids 78-82, 116-120, 190-194, 204-208

N-myristoylation site.

amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 48-59

FIGURE 23

CCCAGAAGTTCAAGGGCCCCCGGCTCCTGCGCTCCTGCCGCCGGGACCCCTCGACCTCCTCA
GAGCAGCCGGCTGCCGCCCGGGGAAGATGCGCAGGAGGAGCCGCCACCGCCTCCTCCTGCTG
CTGCTGCGCTACCTGGTGGTCGCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAA
AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCAA
AGAAGACTGTTTCCTCCAGATTAGAGTGGAAGAACTGGGTCGGAGTGTCTCCTTTGTCTAC
TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTCCAATATCCG
GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTG
AGCAAGGCCAAAACTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT
CCATCATGTGAAGTACCCCTCTTCTGCTCTGAGTGGAACTGTGGTAGAGCTACGATGTCAAGA
CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA
ATCCCAGACTTGGCTCCCAAAGCACCAACAGCTCATACACAATGAATACAAAACCTGGAAC
CTGCAATTTAATACTGTTTCCAAACTGGACACTGGAGAATATTCCTGTGAAGCCCGCAATTC
TGTTGGATATCGCAGGTGTCTTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA
TCATAGCAGCCGTAGTAGTTGTGGCCTTAGTGATTTCGGTTTGIGGCCTTGGTGTATGCTAT
GCTCAGAGGAAAGGCTACTTTTCAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAA
AGCCACGACAATGAGTGAATGTGCAGTGGCTCACGCCGTGAATCCAGCACTTTGGAAGG
CCGCGCGCGCGGATCACGAGGTCAGGAGTTCTAGACCAGTCTGGCCAATATGGTGAAACCC
CATCTCTACTAAATACAAAATTAGCTGGGCATGGTGGCATGTGCCCTGCAGTTCACAGCTGC
TTGGGAGACAGGAGAATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC
CACTGCAGTCCAGCCTGGGTAACAGAGCAAGATTCCATCTCAAAAAATAAAATAAATAAATA
AATAAATACTGGTTTTTACCTGTAGAATTCTTACAATAAATATAGCTTGATATTC

FIGURE 24

MARRSRHRLLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSRLE
WKKLGSRVSFVYYQQTQGD FKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQGQNL EED
TVTLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGQSQT
NSSYTMNTKTGTLQFNTVSKLDTGEYSCEARN SVGYRRCPGKRMQVDDLNISGIIAAVVVVA
LVISVCGLGVCYAQRKGYP SKETSFQKSNSSSKATTMSENVQWLT PVI PALWKAAGGSRGQEF

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 130-144, 238-258

N-glycosylation site.

amino acids 98-102, 187-191, 236-240, 277-281

Casein kinase II phosphorylation site.

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

N-myristoylation site.

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

Amidation site.

amino acids 226-230

FIGURE 25

GACATCGGAGGTGGGCTAGCACTGAAACTGCTTTTCAAGACGAGGAAGAGGAGGAGAAAGAG
 AAAGAAGAGGAAGATGTTGGGCAACATTATTATTAACATGCTCCACAGCCCGGACCCTGGGCAT
 CATGCTGCTATTCTCGAAATACTGAAGAAGCATGGGATTTAAATATTTTACTTCTTAAATAA
 ATGAATTACTCAATCTCCTATGACCATCTATACATACTCCACCTTCAAAAGTACATCAATA
 TTATATCATTAAGGAAATAGTAACCTTCTCTTCTCCAATATGCATGACATTTTGGGCAATG
 CAATTGTGGCACTGGCAGCTTATTTTCAGTGAAGAAAAAATTTGTGGTTCTATGGCAATTCATCA
 TTTGACAATGCAAGCATCTTCTTATCAATCAGCTCCTATTGAACTTACTAGCATCTGACTG
 TGGAACTCCTTAAGGGCCCATACATTTCTGAAGAAGAAAGCTAAGATGAAGGACATGCCACT
 CCGAATTCATGTGCTACTTGGCCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG
 TGGATTGTCCACGGTTATGTACGCTGTGAAATCAGGCCTTGGTTTACACCCAGATCCATTTAT
 ATGGAAGCATCTACAGTGGATTGTAATGATTTAGGTCTTTTAACTTTCCAGCCAGATTGCC
 AGCTAACACACAGATTCTTCTCTACAGACTAACATATTGCAAAAATTGAATACTCCACAG
 ACTTTCCAGTAAACCTTACTGGCCTGGGATTTATCTCAAAACAATTTATCTTTCAGTCACCAAT
 ATTAATGTAAAAAAGATGCCTCAGCTCCTTTCTGTGTACCTAGAGGAAAAACAACITACTGA
 ACTGCCGTGAAAAATGTCTGTCCGAATCAGCAACTACAGAATCTATATTAATCAACAAT
 TGCTTTCTACAATTTCACTCTGGAGCCTTATTGGCCTACAATCTTCTTTCGACTTCATCTC
 AATTCAATAGATTGCAGATGATCAACAGTAAGTGGTTTGTATGCTCTTCCAATCTAGAGAT
 TCTGATGATTGGGGAAAAATCCAATTATCAGAATCAAAGACATGAACCTTTAAGCCTCTTATCA
 ATCTTTCGACGCTGGTTATAGCTGGTATAAACTCACAGAAATACCAGATAACGCCTTGGTT
 GGACTGGAAGCTTAGAAAGCATCTCTTTTACGATAACAGGCTTATTAAAGTACCCCATGT
 TGCTCTTCAAAAAGTTGTAATCTCAAATTTTGGATCTAAATAAAAAATCCTATTAAATAGAA
 TACGAAGGGGTGATTTTAGCAATATGCTACACTTAAAGAGTTGGGGATAAAATAATATGCCT
 GAGCTGATTTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC
 TACTAACCAACCTAGATTGTCTTACATTACCCCAATGCATTTTTCAGACTCCCCAAGCTGG
 AATCACTCATGCTGAACAGCAATGCTCTCAGTGCCCTGTACCATGGTACCATTGAGTCTCTG
 CCAAACCTCAAGGAAATCAGCATACACAGTAACCCCATCAGGTGTGACTGTGTCTATCCGTTG
 GATGAACATGAACAAAAACCAATTCGATTTCATGGAGCCAGATTCACTGTTTTGCGTGGACC
 CACCTGAATTCGAAGGTGAGAATGTTTCGGCAAGTGCATTTTCAGGGACATGATGGAATTTGT
 CTCCTCTTATAGACTCCTCGAGAGCTTTCCTTCTAATCTAAATGTAGAAGCTGGGAGCTATGT
 TTCCTTTCACTGTAGAGCTACTGCAGAACCACAGCCTGAAATCTACTGGATAACACCTTCTG
 GTCAAAACTCTTGCCATAATACCTGACAGACAAGTTCTATGTCCATTCTGAGGGGAACACTA
 GATATAAATGGCGTAACCTCCAAAGAAGGGGGTTTATATACTGTATAGCAACTAACCTAGT
 TGGCGCTGACTTGAAGTCTGTTATGATCAAAGTGGATGGATCTTTTCCACAAGATAACAATG
 GCTCTTTGAATATTAATAAAGAGATATTCAAGGCCAATTCAGTTTTGGTGTCTCTGGAAAGCA
 AGTTCTAAAAATCTCAAATCTAGTGTAAATGGACAGCCTTTGTCAAGACTGAAAAATCTCA
 TGCTGCGCAAGGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACATCTGAATC
 CATCAACTGAGTATAAAATTTGTATTGATATTCCCACTATCTCAGAAAAACAGAAAAAAA
 TGTGTAATGTCAACCACAAAGGTTTGCACCTGATCAAAAAGAGTATGAAAAGATAAATAC
 CACAACACTTATGGCTGTCTTGGAGGCCTTCTGGGGATTATTGGTGTGATATGTCTTATCA
 GCTGCCCTCTCCAGAAATGAACGTGTGATGGTGGACACAGCTATGTGAGGAATTACTTACAG
 AAACCAACCTTTGCATTAGGTGAGCTTTATCTCTCTGTATAAATCTCTGGGAAGCAGGAAA
 AGAAAAAAGTACATCACTGAAAGTAAAGCAACTGTTATAGGTTTACCAACAAATATGCTCTT
AAAAACCAAGGAAACCTATCCAAAAATGAAC

FIGURE 26

MKDMPLRIHVLLGLAITTLVQAVDKKVDPCRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLT
FFARLPANTQIILLQTNNAIKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLE
ENKLTLPKCLSELSNLQELYINHNLSTISPGAFIQLHLLRLHLNSNRLQMINSKWFDA
LPNLLEILMIGENPIIRIKDMNFKPLINLRSLSVIAGINLLEIPDNALVGLLENLESISFYDNRL
IKVPHVALQKVNLKFLDLNKNPINRIRRGDFSNNMLHLKELGINNMPELISIDSLAVDNLPD
LRKIEATNNPRLSYIHPNAPFRLPKLESMLNSNALSALYHGTIESLPNLKEISIHSPNIRC
DCVIRWMNMNKTNI RFMEPDSLFCVDPPEFQGGQNVVRQVHFRDMMELCLPLIAPESFPSNLNV
EAGSYVSFHCRTAEPPQPEIYWITPSGQKLLPNTLTDKFVYVHSEGTLDINGVTFKEGGLYTC
IATNLVAGADLKSVMKIVDGSFPQDNNGSLNLIKIRDIQANSVLVSWKASSKILKSSVKWTAFAV
KTENSHAAQSARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKE
YEKNNTTTLMACLGGLLGIIGVICLISCLSPMNCDDGGHSYVRNYLQKPTFALGELYPLIN
LWEAGKEKSTSLKVKATVIGLPTNMS

Signal sequence:

amino acids 1-22

Transmembrane domain:

amino acids 633-650

N-glycosylation site.

amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583,
608-612, 624-628, 625-629

Casein kinase II phosphorylation site.

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

Tyrosine kinase phosphorylation site.

amino acids 570-579

N-myristoylation site.

amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443,
491-497, 492-498, 634-640, 702-708

Cell attachment sequence.

amino acids 277-280

FIGURE 27

GCCCGGGA CTGGCGCAAGGTGCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTTAG
CTGCAGCCTTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC
CAGCCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCAGCATATGAATCT
GGTAGACCTGTGGTTAACCGTTCCCTCTCCATGTGTCTCCTCTACAAAGTTTTGTTCTTA
TGATACTGTGCTTTATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCTCTGGG
GGTTTAAATGTCACTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCTCTCTGA
AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGG
ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT
GCCTTCAAAGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAG
TGTGCACAAAAATGCCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACACCCCTGGC
ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCAC
AACGTGATCTGTAAAAAGTCCGTGTTGGATGAACATGCTGGCAGACCATTCCTCAATGCTGC
CAACGACGCTGACCTTGTAACTCCCTAAAAAACTACCGATTATGCCATGCTGGTCACCA
TGTTTGCTGGTTCACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAG
GATGCCCGGAGACACCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGA
ACCTGATGATATTAGCACTGTGGTATTAGGTGTCCAACTGACTGTCTATTGAGAAAGAAAGAAA
GTAGTTTGCGATTGCAGTAGAAATAAGTGGTTTACTTCTCCCATCCATTGTAAACATTTGAA
ACTTTGTATTTTCAGTTTTTTTTGAATTATGCCACTGCTGAACTTTTAAACAACTACAAACA
TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCGATGGTATATTTCTGAGT
AAGCTACTATCTGAACATTAGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTT
AATTTAAAGCAAATAAAGCTTAACTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAACA

FIGURE 28

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMC PKGCLCSSSGGLNVTC SNANLKEIPRDL
PPETVLLYLD SNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGV AETLQTLDLSDNR
IQSVHKNAFNNL KARARIANNPWHCDCTLQQVLRSMASNHETAHNVICKTSVLDEHAGRPFL
NAANDADLCNL PKKTTDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYL KSLPSRQKK
ADEPDDISTVV

Signal sequence:

amino acids 1-33

Transmembrane domain:

amino acids 205-220

N-glycosylation site.

amino acids 47-51, 94-98

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 199-203

Casein kinase II phosphorylation site.

amino acids 162-166, 175-179

N-myristoylation site.

amino acids 37-43, 45-51, 110-116

FIGURE 29

ACCGAGCCGAGCGGACCGAAGGCGCGCCCGAGATGCAGGTGAGCAAGAGGATGCTGGCGGGG
GGCGTGAGGAGCATGCCAGCCCCCTCCTGGCTGTGGCAGCCCCATCCTCCTGCTGGTGCT
GGGCTCAGTGCTGTGAGGCTCGGCCACGGGCTGCCGCCCGCTGCCAGTGCTCCGCCCCAGG
ACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAGTCCCCGAGGGCATCCCAACCGAG
ACGCGCTCTGCTGGACCTAGGCAAGAACCAGCATCAAAACGCTCAACCAGGACGAGTTCCGCCAG
CTTCCCGCACTTGGAGGAGCTGGAGCTCAACGAGAAACATCGTGAGCGCCGTGGAGCCCGCGG
CCTTCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGCAGCAACCGCTGAAGTCTCATC
CCGCTAGGCGTCTTCACTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAAACAAGAT
CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG
ACAATGACCTCGTCTACATCTCTCACCGCGCCTTCAGCGGCTCAACAGCCTGGAGCAGCTG
ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCCACCTGCACGGCT
CATGTCCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC
TGTACCGACTCAAGTCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCCAAC
TGCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCC
CTACCTGGCCGTCCGCCACCTAGTCTATCTCCGCTTCTCAACCTCTCCTACAACCCCATCA
GCACCATTTGAGGGCTCCATGTTGTCATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGGC
GGGAGCTGGCCGTGGTGGAGCCCTATGCCCTCCGCGGCTCAACTACCTGCGCGTGTCTCAA
TGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG
AGACACTCATCCTGGACTCCAACCGCTGGCCTGCGACTGTGCGCTCCTGTGGGTGTTCCGG
CGCCGCTGGCGCTCAACTCAACCGGCAGCAGCCACGTGCGGCCACGCCCGAGTTTGTCCA
GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACTGCGCCGCG
CCCGCATCCGGGACCCGAAGGCCAGCAGGTGTTTGTGGACGAGGGCCACACGGTGCAGTTT
GTGTGCCGGGCCGATGGCGACCCGCCGCCGCCATCCTCTGGCTCTCACCCCGAAAGCACT
GGTCTCAGCCAAGAGCAATGGCGGCTCACAGTCTTCCCTGATGGCAGCTGGAGGTGCGCT
ACGCCAGGTACAGGACCAACGGCACGTACCTGTGCATCGCGGCCAACCGGGCGGCAACGAC
TCCATGCCCGCCCACTGTCATGTGCGCAGCTACTCGCCGACTGGCCCCATCAGCCCAACAA
GACCTTCGCTTTTCATCTCCAACAGCCGGGCGAGGGAGAGGGCCAACAGCACCCGCGCCACTG
TGCCTTTCCCTTTCGACATCAAGACCCCTCATCATCGCCACACCATGGGCTTTCATCTCTTTC
CTGGGCGTCTCTCTTCTGCTTGGTGCTGCTGTTTCTCTGCGGCGGGGCAAGGGCAACAC
AAAGCACAAACATCGAGATCGAGTATGTGCCCGAAAGTCGGAACGAGGATCAGCTCCGCCG
ACGCGCCCCGAAGTTCAACATGAAGATGATATGAGGCCGGGGCGGGGGGCGAGGACCCCCG
GGCGGCGGGCAGGGGAAGGGGCTGGTCGCCACCTGCTCACTCTCAGTCTTCCACCTC
CTCCTACCCCTTCTACACGTTCTCTTCTCCTCCGCTCCGCTCCCTGCTGCCCCCCG
CCAGCCTCACCACTGCGCTCTTCTACCAGGACCTCAGAAGCCGAGACCTGGGGACCCCA
CCTACAGGGGCAATTGACAGACTGGAGTTGAAAGCCGACGAACCCGACACGCGGAGAGTCA
ATAATTCAATAAAAAAGTTACGAACCTTCTCTGTAACTTGGGTTTCAATAATTATGGATTTT
TATGAAACTTGAATAAATAAAGAGAAAAAACTAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MQVSKRMLAGGVSRMSPSELLACWQPILLVLGSLVSGSATGCPPRCECSAQDRAVLCHRKCF
VAVPEGIPTETRLDLGKNRIKTLNQDEFASFPHLEELNENIVSAVEPGAFNNLNLRLTL
GLRSNRKLKLIPLGVFTGLSNLTQDISENKIVILLDYMFDLYNLKSLEVGDNLDLVYISHRA
FSGLNSLEQLTLEKCNLTISIPTALSHLHGLIVLRRLNLNINAIIRDYSFKRLYRLKVLEISH
WPYLDTMTNPNCLYGLNLTSLSIHCNLTAVPYLAVRHLVYLRFLNLNSYNPISTIEGSMLEH
LRLQETQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTTLLESVPHSVGNLETLILDSNPLA
CDCRLLWVFRRRWRNLNFRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV
FVDEGHTVQFVCRADGDDPPAILWLSPRKHLVSAKSNGRLTVPDPGTLEVRYAQVQDNGTYL
CIAANAGGNDMPAHLHVRSYSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI
IATTMGFISFLGVVLFCLVLVLFWSRGKGNTKHNIEYVPRKSDAGISSADAPRKFNMKMI

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 556-578

N-glycosylation site.

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345,
492-496, 505-509, 526-530, 542-546

Casein kinase II phosphorylation site.

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353,
607-611

Tyrosine kinase phosphorylation site.

amino acids 590-598

N-myristoylation site.

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143,
262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 32-43

FIGURE 31

CCCACGCGTCCGCACCTCGGCCCGGGCTCCGAAGCGGCTCGGGGCGCCCTTTTCGGTCAAC
ATCGTAGTCCACCCCTCCCATCCCCAGCCCCGGGGATTAGGCTCGCCAGCGCCAGCC
AGGGAGCCGGCCGGGAAGCGCGATGGGGGCCCCAGCCGCTCGTCTCTGCTCCTGCTCCTGC
TGTTTCGCTGCTGCTGGGCGCCCGCGGGGCCAACCTCTCCAGGACGACAGCCAGCCCTGG
ACATCTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGTCAAGTGCCAGTGAAAGATCA
CGAGGACTCATCCCTGCAATGGTCTAACCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGA
GAGCCCTTCGAGATAATCGAATTACGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGC
ATCAGCAATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGT
GCGAACTGCCAAGTCCCTCGTCACTGTGCTAGGAATCCACAGAAGCCCATCATCACTGGTT
ATAAATCTTCATTACGGGAAAAAGACACAGCCACCTTAACTGTCACTCTCTGGGAGCAAG
CCTGCAGCCCGGCTCACCTGGAGAAAGGGTGACCAAGAACTCCACGGAGAACCAACCCGCAT
ACAGGAAGATCCCAATGGTAAACCTTCACTGTGAGCAGCTCGGTGACATTCAGGTTACCC
GGGAGGATGATGGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGAC
AGATCCACCTCTCAACGCATTGAAGTTTTATACACCAACTGCGATGATTAGGCCAGACCC
TCCCCATCTCTGTGAGGGCCAGAAGCTGTTGCTACACTGTGAGGGTCGCGGCAATCCAGTCC
CCCAGCAGTACCTATGGGAGAAGGAGGGCAGTGTGCCACCCCTGAAGATGACCCAGGAGAGT
GCCCTGATCTTCCCTTTCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAG
CAACATGGGCAGCTACAAGGCCTACTACACCCCTCAATGTTAATGACCCAGTCCGGTGCCCT
CCTCCTCCAGCACTACCACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCCTGCTG
CTCATCATGCTCATCTTCTTGGCCACTACTTGATCCGGCACAAAGGAACCTACCTGACACA
TGAGGCAAAAGGCTCCGACGATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCG
GGCAGTCAGGAGGGGACGACAAGAAGGAATATTTTCATCTAGAGGCGCCTGCCCACTTCTCTGC
GCCCCCAGGGGCCCTGTGGGGACTGCTGGGGCCGTACCAACCCGGACTTGACAGAGCAA
CCGACGGGCCGCCCCCTCCGCTTGCTCCCAGCCCCACCCACCCCTGTACAGAATGTCTGC
TTTGGGTGCGGTTTTGTACTCGGTTTGGAATGGGAGGGAGGAGGGCGGGGGAGGGGAGGG
TTGCCCTCAGCCCTTTCGTTGGCTTCTCTGCATTGGGTTATTATTATTTTGTAAACATCC
CAAATCAAATCTGTCTCCAGGCTGGAGAGGCAGGAGCCCTGGGGTGAGAAAAGCAAAAAACA
AACAAAAACA

FIGURE 32

MGAPAASLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTIVLKCQVKDHEDSSLQW
SNPAQQTLYFGEKRALRDNR IQLVTSTPHELSSISINVALADEGEYTC SIFTMPVRTAKSLV
TVLGIPQKPI IITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRI QEDPNGK
TFTVSSSVTFQVTRREDDGASIVCSVNHESLKGADRSTSQR IEVLYTPTAMIRDP PHPPREGQ
KLL LHCEGRGNPVPQQYLWEKEGSPPLKMTQESALIFPFLNKSDSGTYGCTATSNMG SYKA
YYTLNVNDPSPVPSSSSSTYHAIIGGIVAFIVFLLLIMLIFLGHYLRHKGTYLTHEAKGSDD
APDADTAIINAEGGQSGGDDKKEYFI

Signal sequence:

amino acids 1-20

Transmembrane domain:

amino acids 331-352

N-glycosylation site.

amino acids 25-29, 290-294

Casein kinase II phosphorylation site.

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

N-myristoylation site.

amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304,
306-310, 334-340, 360-364, 385-389, 386-390

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 33

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCAAACCCCTTTTCTTCTCCTTTCTCTGG
CTTCGGACATTGGAGCACTAAATGAACCTGAATTTGTCTGTGTGGCGAGCAGGATGGTTCGCTG
TTACTTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAAATGCTGCTTTGGATTCTGTT
GCTGGAGACGTCTCTTTGTTTTGCCCGCTGGAACGTTACAGGGGACGTTTGCAAAAGAGAAGA
TCTGTTCTGCAATGAGATAGAAGGGACCTACACGTAGACTGTGAAAAAAGGGCTTCACA
AGTCTGCGAGCGTTTCACTGCCCCGACTTCCCAGTTTACCATTATTTCTGCATGGCAATTC
CCTCACTCGACTTTTCCCTAATGAGTTCCGCTAACTTTTATATACTGGGTTAGTTTGACATGG
AAAACAATGGCTTGCATGAATCGTTCCGGGGGCTTTTCTGGGCTGACGCTGGTGAAAGG
CTGCACATTCAACACAAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA
TCTGGAATATCTCCAGGCTGATTTTAATTTATTACGAGATATAGACCCGGGGGCTTCCAGG
ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACATCTCATCAGACCCCTACCTGCCAAC
GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC
CTATGAGGAGGTCTTGGAGCAAACTCCCTGGTATTGCGGAGATCTGTAGAGGATAACCCCTT
GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAAACATTTCCCAAGAATGCC
CTGATCGGCCCGAGTGGTCTGCGAAGCCCCACAGACTGCGAGGTAAAGACCTCAATGAAC
CACGGAACAGGACTTTGTCTCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGCGCCCCCTG
CCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCAAGAG
GATCATGCCACACAGGGTCTGCTCCAAACCGGAGGTACAAAGATCCCAGGCAACTGGCAGAT
CAAAATCAGACCCACAGCAGCGATAGCGAGGGTAGCTCCAGGAACAAACCTTAGCTAAC
GTTTACCCTGCCCTGGGGGCTGCAGCTGCGACCACATCCAGGGTCTGGGTTTAAAGATGAAC
TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTCTAACGTGCAGGA
GCTTTTCTACGAGATAACAAGATCCACAGCATCCGAAAACTCGCACTTTGTGGATTACAAGA
ACCTCATTCTGTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACCAACACTTTCAAG
AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGA
GAAATTCTCGGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA
TCTTCCCGGGCACTTTCAATGCCATGCCAAACTGAGGATCCTCATTCTCAACAACAACCTG
CTGAGGTCCCTGCCTGTGGACGTGTTCTGCTGGGGTCTCGCTCTCTAAACTCAGCCTGGCAAA
CAATTACTTCATGTACTTCCCGTGGCAGGGGTGCTGGACCACTTAACCTCCATCATCTCAGA
TAGACCTCCACGGAACCCCTGGGAGTGCTCTTGCACAATTGTGCTTTCAAGCAGTGGGCA
GAACGCTTGGGTTCCGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGCTGAACTTCTT
TAGAAAGGATTTCACTGCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT
CGCCCACGTAACTTCTGCACAGTAAAAACAGCACTGGGTTGGCGGAGACCCGGGACGCATCC
AACTCTACCTAGACACGAGCAGGGTGTCCATCTCGGTGTTGGTCCCGGACTGTGCTGGT
GTTTGTACCTCCGCTTACCCTGGTGGGCTGCTCGTGTTTATCTGAGGAACCGAAAGC
GGTCCAAGAGACGAGATGCCAACTCCTCCGCGTCCGAGATTAATTTCCCTACAGACAGTCTGT
GACTCTTCTTACTTGGCACAATGGGCCTTACAACGCAGATGGGGCCCCACAGAGGTGATGACTG
TGGCTCTCACTCGCTCTCAGACTAAGACCCCAACCCCAATAGGGGAGGCGAGAGGAAGCG
ATACATCCTTCCCACCGCAGGCACCCCGGGGCTGGAGGGCGTGTACCCAAATCCCCGCG
CCATCAGCTGTGATGGGCATAAGTAGATAAAATAACTGTGAGCTCGCACAACCGAAAGGGCT
GACCCCTACTTTAGCTCCTCCTTGAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA
GCCAGCTCGCTCTTTGCTGAGAGCCCCCTTTGACAGAAAGCCAGCACGACCTGCTGGAAG
AACTGACAGTGCCCTCGCCCTCGGCCCCGGGGCCTGTGGGGTTGGATGCCCGGGTTCTATAC
ATATATACATATACCATCTATATAGAGATAGATATCATTTTTCCTGTTGGATTAG
CCCGTGATGGTCCCTGTTGGCTACGCAGGATGGGCAGTTGCACGAAGGCATGAATGTAT
TGTAATAAGTAACCTTGACTTCTGAC

FIGURE 34

MLLWILLLETSLCFAAGNVTGDVCKEIKCSCNEIEGDLHVDCEKKGFTSLQRFTAPTQSQFYH
LFLHGNSLTRLPFNEFANFYNAVSLHMENGLHEIVPGAFGLGLVVKRLHINNKKIKSFRKQ
TFLGLDDLEYLQADFNLRLRIDPGAFQDLNKLEVLILNDNLISTLPANVFQYVPIITHLDLRG
NRLKTLTPYBEVLBQIPGIABIILEDNPWDCTCDLLSLKWELENI PKNALIGRVVCEAPTRLQ
GKDLNETTEQDLCPLKNRVDSSLPAPPAQEETFAPGPLPTPFKTNQEDHATPGSAPNGGTK
IPGNWQIKIRPTAAIATGSSRNKPLANS LPCPGGCS CDHIPGSGLKMN CNNRNVSS LADLKP
KLSNVQELFLRDNKIHSIRKSHFVDYKNLILDLGNNNIATVENNTFKNLLDLRWLYMDSNY
LDTLSREKFAGLQNLVYLNVEYNAIQLILPGTFNAMPKLRILILNNNLLRSLPVDVFAGVSL
SKLSLHNNYFMYLPVAGVLDQLTSIIQIDLHGNPWECSCTIVPFKQWAERLGSEVLMSDLKC
ETPVNFFRKDFMLLSNDEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVL
VPGLLLVFVTS AFTTVVGM LVFILRNKRKSKRRDANSSASEINSLQTVCDDSSYWHNGPYNADG
AHRVDCGSHSLSD

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 618-638

N-glycosylation site.

amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 122-126, 646-650

Casein kinase II phosphorylation site.

amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577,
608-612, 657-661, 666-670, 693-697

N-myristoylation site.

amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349,
354-360, 465-471, 493-499, 598-604, 603-609

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 337-348

Cellulose

AGTGCAGCTGCGTCCCCGTGTAACCCGGCGCCAGCTGTGTTCCTGACCCAGAAATAACTCAGGGC
TGACCCGGGGCTCGGACGCGCTCCGACACATTTCTCTGTCGGCCCTAAGGGAAACTGTGGTC
CGCTGGGGCTCGGGGGGATTCTTGCAGTTGGGGGGTCCGTGCGGAGACAGGGCGGAGGGG
AAGGGAGGGGAAACCGGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC
AGCTCTGCGTCTCTCGAGCGGGAAGATCAAGTTTGGGAGCAGCTCTGCGTGGCGGGCTCAG
AGAATGAGGGCGGGCTTCGCCCTGTGTCCTCTCTGGCAGGCGCTCTGGCCGCGGGCGGGCGG
CGGCGAACAACCCCACTGCGCAGCGTGTCTGGCTGTCTCGGCTCTGGGGGCGCTGCTACAGCCTG
ACCACGCTACCATTGAAGCGCAGGCGCGCGAGGAGCCTGCATCTCTCGAGGTGGGGCGCTC
AGCACCGTGTGCTGGGGCGCCGAGCTGCGCGCTGTGCTCTGCGCTCTCTCGGGACGGGCCAGG
CCCGGAGAGGGGGCTCAAAGACCTGCTGTTCTGGGTGCATCGGAGCGAGCGGCTTCCCACT
GCACCTGTGAGAAAGACGACTTTTGGGGTTGCTCTCGGTGTCTTCCGACCCCGCGGCTCTC
GAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCAACGCTCTCTGACC CGCGGAGATGCGC
GGTACTCTCAGGCGACCGGTGGGGTGCAGCGCCGAGCGTGGAAAGAGATGCGATGCCACTGCG
GCGCAACGGCTACTGTGTCAAGTACCAAGTTTGAGGCTTTGTGTCTCTGCGCGCGCCGCGG
GCGGCTCTAACTTGAGCTATCGCGCGCCCTTCCAGCTGCACAGCGCGCTCTGGAATTGCA
TCCCACTGGGACAGGTTGAGTGTGCTCTGCGGGGACAGCTCCCGATCTCAGTTACTTGCT
CTCGGGAGCAAACTCGGCGCTCGCTGGGACAAACTCTCGGCGAGTGTGTGTCTCTGCC
GGGAGGTACTCTCGTGTGTGGCAAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGG
CTTTGGCTCGGAATTGTCTACGGGCTTCAGCTGGGGGAAGGACGGCCGCTCTGTGTGACCA
GTGGGGAAGGACAGCGACCTCTGGGGGACCGGGTGCCACAGGCGCCCGCGGCACT
GCAACACCGCCCGTGCAGCAGAAACATGCCAATCAGGGTCGACAGAGAAGCTGGGAGAGAC
ACCACTTGTCCCTGAACAAGACAATCAGTAACTATCTTCTAGATTCTCGATGGGAGT
CACAGAGCAGATGTCTACCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCACC
CATCAGGAGCGGTGATTTCAAAGTTAAATCTACAGACTCTCTCGGCACTCTCAGGCTTT
CGACTCTCTCTGCGGTGTCTTCAATTTGTGAGCAGACAGTAGTAGTGTGTGTGATCT
TGACCATGACAGTACTTGGGGCTTGTCAAGCTCTGCTTACGAAAGCCCTCTTCCAGCCA
AGGAAGGAGTCTATGGGCGCGCGGGCTGGAGTGTATCTGAGCCCGCTGCTTTGGGCTC
CAGTTCTGCACATTGCACAAACAATGGGGTGAAGAGTCGGGACTGTGATCTGCGGACAGAG
CAGAGGGTGCTTCTGTGCGGAGTCCCCCTTTGGCTCTAGTGATGCATAGGGAACAGGGGA
CATGGGCACTCTGTGAACAGCTTTTCACTTTTGATGAACCGGGGAACCAAGAGGAACCTAG
TTGTGTAAGTACAATTTCTGCAGAAATCCCCCTTCTCAAATTCCTTTACTCCAAGTACG
GAGCTAAATCAGAAGTCGACACTCTCTCCCTGATGATAGAGGAAGTGAAGAGTGCCTTTAGGA
TGGTGAATCTGGGGGACCGGGTAGTGTCTGGGAGAGATATTTTCTTGTATTTCGAGAA
TTTGGAGAAGTGATTGAACCTTTCAAGCATTGGAAACAATAAGAACCAATATAATTTACA
TTAAATAATAATTTCTACCAAAATGGAAGGAATGTCTATGTGTTCTCAGGTCAGGAGTAT
ATTGGTTGCAAAATCCGAGGGAATAAAATAAAAAATAAAGGATTTGGTGTAT

FIGURE 36

MRPAFALCLLWQALWPGPGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALS
TVRAGAE LRAVLALLRAGPGPGGGSKD L LFWVALERRRSHCTLENEPLRGFSWLSSDPGGLE
SDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLC PAPRPGA
ASNL SYRAFFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADRI GARWDKLSGDVLCPCPG
RYLRAGKCAELPNCLDDLGGFACCATGFELGKDGRSCVTS GEGQPTLGGTGVPTRRPPATA
TSPVPQRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP
SGSVISKFNSTTSSATPQAPDSSSAVVFI FVSTAVVVLVILTMTVLGLVKLCFHESFSSQER
KESMGPPGLES DPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 189-193, 381-385

Glycosaminoglycan attachment site.

amino acids 289-293

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 98-102, 434-438

Casein kinase II phosphorylation site.

amino acids 275-279, 288-292, 342-346, 445-449

N-myristoylation site.

amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157,
185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469,
477-483

Aspartic acid and asparagine hydroxylation site.

amino acids 262-274

FIGURE 37

CGGACGCGTGGGATTAGCAGTGGCCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAACTAAG
CGTCGAGTCAGACGGCACCATAATCGCCTTTAAAGTGCCTCCGCCCTGCCGGCCGCGTATC
CCCCGGCTACCTGGGCGCCCCGCGCGGTGCGCGCTGAGAGGGAGCGCGCGGGCAGCCGA
GCGCCGCTGTGAGCCAGCGCTGCTGCCAGTGTGAGCGCGCGGTGTGAGCGCGGTGGGTGCGGA
GGGGCGTGTGTGCCGCGCGCGCGCCGTGGGGTGCAAAACCCGAGCGTCTACGCTGCCATGA
GGGGCGCGAACGCGCTGGGCGCCACTCTGCCTGCTGCTGGCTGCGGCCACCCAGCTCTCGCG
CAGCAGTCCCCAGAGAGACCTGTTTTCACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT
TATTGGCAGTGAAGGTTTTCCTGGAGTGTACCCTCAAATAGCAAATGTACTTGGAAATCA
CAGTTCCCGAAGGAAAGTAGTCGTTCTCAATTTCCGATTTCATAGACCTCGAGAGTGACAAC
CTGTGCCGTATGACTTTGTGGATGTGTACAATGGCCATGCCAATGGCCAGCGCATTGGCCG
CTTCTGTGGCACTTTCCGGCCTGGAGCCCTTGTGTCCAGTGGCAACAGATGATGGTGCAGA
TGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATGTTCTCCGCTGCTGAACCA
AACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGACCTTCCGGCTCTTTTAAAC
CCCCAACTGGCCAGACCGGATTAACCTGACAGGAGTCACTTGTGTGTGGCACTTTAGCCC
CAAAGAATCAGCTTATAGAATTAAGATTGAGAAGTTGATGTGGAGCGGAGATAACTACTGC
CGATATGATTATGTGGCTGTGTTTAAATGGCGGGGAAGTCAACGATGCTAGAGAATTGGAAA
GTATTGTGGTGATAGTCCACCTGCGCCAATTGTGTCTGAGAGAATGAACTTCTTATTAGT
TTTTATCAGACTTAAGTTTAACTGCAGATGGGTTTATTTGGTCACTACATATTCAGGCCAAAA
AAACTGCCCTACAACACAGAACAGCCTGTCAACACCACATTCCTGTAAACACGGGTTTAAA
ACCCACCGTGGCCTTGTGTCAACAAAGTGTAGACGGACGGGGACTCTGGAGGGCAATTATT
GTTCAAGTGACTTTGTATTAGCCGGCACTGTATCAACAACCATCACTCGCGATGGGAGTTTG
CAGCCACAGTCTCGATCATCAACATCTACAAAGAGGGAATTTGGCGATTTCAGCAGGGCGG
CAAGAACATGAGTGCCAGGCTGACTGTGCTGTGCAAGCAGTGCCTCTCCTCAGAAGAGGTC
TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGGCGAGGCAAAATCATGCCAAACAGC
TTTATCATGATGTTCAAGACCAAGAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATG
TTAAACAGTGAAGTGTGCTCAATTTAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTT
TCTCAGTAGAAAAAAATACCTTATAAAATTACATATTTGAAAGAGGATTCCGAAAGATGG
GACTGGTTGACTCTTACATGATGGAGGTATGAGGCCTCCGAGATAGCTGAGGGAAGTTCCT
TGCTGCTGTCTCAGAGGAGCAGCTATCTGATTGGAACCTGCCACTTGTGCGGTGATAGGA
AGCTAAAAAGTGTCAAGCGTTGACAGCTTGGAAAGCGTTTATTTATACATCTCTGTAAAAGGAT
ATTTTAGAATTGAGTTGTGTGAAGATGTCAAAAAAGATTTTAGAAGTGCAATATTTATAGT
GTTATTTGTTTCACTTCAAGCCTTTGCCCTGAGGTGTTACAATCTGTCTTGTGCGTTCCTTA
AATCAATGCTTAATAAAATATTTTAAAGGAAAAA

FIGURE 38

M R G A N A W A P L C L L L A A A T Q L S R Q Q S P E R P V F T C G G I L T G E S G F I G S E G F P G V Y P P N S K C T W K
I T V P E G K V V V L N F R F I D L E S D N L C R Y D F V D V Y N G H A N G Q R I G R F C G T F R P G A L V S S G N K M M V
Q M I S D A N T A G N G F M A M F S A A E P N E R G D Q Y C G G L L D R P S G S F K T P N W P D R D Y P A G V T C V W H I V
A P K N Q L I E L K F E K F D V E R D N Y C R Y D Y V A V F N G G E V N D A R R I G K Y C G D S P P A P I V S E R N E L L I
Q F L S D L S L T A D G F I G H Y I F R P K K L P T T T E Q P V T T T F P V T T G L K P T V A L C Q Q C R R T G T L E G N
Y C S S D F V L A G T V I T T I T R D G S L H A T V S I I N I Y K E G N L A I Q Q A G K N M S A R L T V V C K Q C P L L R R
G L N Y I I M G Q V G E D G R G K I M P N S F I M M F K T K N Q K L L D A L K N K Q C

Signal sequence:

amino acids 1-23

N-glycosylation site.

amino acids 355-359

Casein kinase II phosphorylation site.

amino acids 64-68, 142-146, 274-278

Tyrosine kinase phosphorylation site.

amino acids 199-208

N-myristoylation site.

amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295,
305-311, 309-315, 320-326, 330-336

Cell attachment sequence.

amino acids 149-152

FIGURE 39

CGGACGCGTGGGCGGACGCGTGGGCGGCCACGGCGCCCGCGGGCTGGGGCGGTGCGCTTCTT
CCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAGATGGCCCATGGCCCCGAAGG
GCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCTCAACCTCCAGGACCTATCTGG
CTCCAGCCCTCTCCACCTCCCCAGTCTTCTCCCCGCTCAGCCCCATCCGTGTATACCTG
CCGGGGACTGTTTGACAGCTTTAAACAAGGGCCTGGAGAGAACCATCCGGGACAACTTTGGAG
GTGGAAACACTGCCTGGGAGGAAGAAATTTGTCCAATAACAAGACAGTGAGACCCGCCTG
GTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT
GAGTGAGGAGCTGGTGGAGAGCTGGTGGTTTCAAGCAGCAGGAGGCCCGGACCTCTTCC
AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCGCAGGCACCTTCGGGCCCTCCTGC
CTTCCCTGTCTGGGGGAACAGAGAGGCCCTGCGTGGCTACGGGCAGTGTGAAGGAGAAGG
GACACGAGGGGGCAGCGGCCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCCTGTGGCC
AGTGTGGCTTTGGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTCCGGCTTGT
TTTGGCCCTGTGCCCCGATGCTCAGGACCTGAGGAATCAAACCTGTTTGAAGTGAAGAAGGG
CTGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAAC
GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG
GCCTGCCTAGGTGCATGGGGGCAGGGCCAGGTGCTGTAAGAAGTGTAGCCCTGGCTATCA
GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGA
ACAAGCAGTGTGAAAACACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG
ATGGAAGGCATCTGTGTGAAGGAGCAGATCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC
AGAAGACGAGTTGGTGGTGTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCA
CGCTGGCTGCTAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATG
ACTGGCTACTGTTGTGACAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATA
ATCGCGGCCACCACTGTAGACCTCCTCCACCCACGCTGCCCCCAGAGCTTGGGCTGCCC
TCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTG
CCTTACAGAGCAGCCAGGTACCCAGGCCCGGGCAGACAAGGCCCTGGGGTAAAAAGTAGC
CCTGAAGGTGGATACCATGAGCTCTTCACTGGCGGGGACTGGCAGGCTTCAACATGTGTGA
ATTTCAAAGTTTTTCTTAATGGTGGCTGCTAGAGCTTTGGCCCCCTGCTTAGGATTAGGTG
GTCCTCAGAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGCTGCCAGTTCTGT
TCTGTGTTCCACCACATCCCCACACCCCATTGCCACTTATTTATTCATCTCAGGAAATAAAGA
AAGGCTCTGGAAGTTAAAAA

CGTGGTAAAGT

FIGURE 40

MAPWPPKGLVPAVLWGLSLFLNLP GPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLER
TIRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKQ
QEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGGQCEGEGTRGGSGHCDQAG
YGGEACGQCGLGYFEAERNASHLVCSACFGPCARCSGPPEESNCLQCKKGWALHHLKCVDIDE
CGTEGANCGADQFCVNTESYSYECRDCAKACLGC MGAGPGRCKKCSPGYQQVGSKCLDVDECE
TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDLVVLQQMFFG
IIICALATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

Signal sequence:

amino acids 1-29

Transmembrane domain:

amino acids 372-395

N-glycosylation site.

amino acids 79-83, 205-209

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 290-294

Casein kinase II phosphorylation site.

amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

N-myristoylation site.

amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179,
177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289,
326-332, 372-378, 395-401

Aspartic acid and asparagine hydroxylation site.

amino acids 321-333

EGF-like domain cysteine pattern signature.

amino acids 181-193

FIGURE 41

TGAGACCCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA
GCACCATGCAGCCCCCTGTGGCTCTGCTGGGCACTCTGGGTGTTGCCCTTGGCCAGCCCCGGG
GCCGCCCTGACCGGGGAGCAGCTCCTGGGCAGCCTGCTGCGGCAGCTGCAGCTCAAAGAGGT
GCCCACCCTGGACAGGGCCGACATGGAGGAGCTGGTTCATCCCCACCACGTGAGGGCCAGT
ACGTGGCCCTGCTGCAGCGCAGCCACGGGGACCGCTCCCGCGAAAGAGGTTTCCAGCCAGAGC
TTCGAGAGGTGGCCGGCAGGTTTCTGGCGTTGGAGGCCAGCACACCTGCTGGTGTTCGG
CATGGAGCAGCGGCTGCCGCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTCTTCCAGG
AGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGGCGGCTGTCCCCGCGCAGCGCCCCGGGCC
CGGGTGACCGTCGAGTGGCTGCGCGTCCGCGACGACGGCTCCAACCGCACCTCCCTCATCGA
CTCCAGGCTGGTGTCCGTCCACGAGAGCGGCTGGAAGGCCCTTCGACGTGACCGAGGCCGTGA
ACTTCTGGCAGCAGCTGAGCCGGCCCCGGCAGCCGCTGCTGCTACAGGTGTCGGTGCAGAGG
GAGCATCTGGGCCCGCTGGCGTCCGGCGCCACAAGTGGTCCGCTTTGCTCTCGCAGGGGGC
GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCTGGACCTTGGGGACTATGGAG
CTCAGGGCGACTGTGACCTGAAGCACCAATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG
ATGTACATTGACCTGCAGGGGATGAAGTGGGCCGAGAACTGGTGCTGGAGCCCCGGGCTT
CCTGGCTTATGAGTGTGTGGGCACCTGCCGGCAGCCCCGGAGGCCCTGGCCTTCAAGTGGC
CGTTTCTGGGGCCTCGACAGTGCATCGCCTCGGAGACTGACTCGCTGCCCATGATCGTCAGC
ATCAAGGAGGGAGGCAGGACCAGGCCCCAGGTGGTCAGCCTGCCAACATGAGGGTGCAGAA
GTGCAGCTGTGCCTCGGATGCTGCGCTCGTGCCAAGGAGGCTCCAGCCATTAGGCGCCTAGTG
TAGCCATCGAGGGACTTGACTTGTGTGTGTTTCTGAAGTGTTCGAGGGTACCAGGAGAGCTG
GCGATGACTGAACTGCTGATGGACAAATGCTCTGTGCTCTCTAGTGAGCCCTGAATTTGCTT
CCTCTGACAAGTTACCTCACCTAATTTTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGA
GAGCCCTTGCTCAGTTTTCTATTCTTATTATTACTGCACTATATTCTAAGCACTTACAT
GTGGAGATACTGTAACCTGAGGGCAGAAAGCCANTGTGTGTCATTTGTTTACTTGTCTGTCTAC
TGGATCTGGGCTAAAGTCTCTCACCACCCTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT
TGTGCATCCCCAATCCAGATAATAAAGACTTTGTAAACATGAATAAACACATTTTATTCT
AAAA

FIGURE 42

MQPLWLCWALWVLPASPGAAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQYV
ALLQRSHGDRSRGKRFSQSFRVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRRLFQEP
VPKAALHRHGRILSPRSARARVTVEWLRVRDDGSNRTSLIDSRLSVHESGWKAFDVTEAVNF
WQQLSRPRQPLLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTLDLGDYGAQ
GDGCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCTCRQPPEALAFKWPF
LGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKCSASD GALVPRRLQP

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 158-162

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 76-80

Casein kinase II phosphorylation site.

amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

N-myristoylation site.

amino acids 19-25, 156-162, 225-231, 260-266, 274-280

Amidation site.

amino acids 74-78

TGF-beta family signature.

amino acids 282-298

FIGURE 43

GTCTGTTCCAGGAGTCCTTCGGCGGCTGTTGTGTGTCAGTGGCCTGATCGCGATG**GGG**GACAAA
GGCGCAAGTCGAGAGGAAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG
CATTGGGCAGTGTACAGTGCACCTCTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT
GTGAAGTTGTCTGTGCCTACTCGGGCTTTTCTTCTCCCGTGTGGAGTGAAGTTTGACCA
AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG
TGACCTTCTTGCCAACCTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC
ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCCTCCTCTGCCACCATTGGGAACCGGG
CAGTGCTGACATGCTCAGAACAAGATGGTTCCTCCACCTTCTGAATACACCTGGTTCAAAGAT
GGGATAGTGATGCCATACGAATCCCAAAGCACCCGTCGCTTCAGCAACTCTTCCTATGTCTCT
GAATCCCAACAGGAGAGCTGGTCTTTGATCCCTGTGAGCTCTGATACTGGAGAATACA
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT
GTGGAGCGGAATGTGGGGTTCATCGTGGCAGCGCTCTTGTAACTTGAATTCCTCTGGGAAT
CTTGGTTTTTGGCATCTGGTTTGCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA
CTTCGAGTAAGAAGGTGATTACAGCCAGCCTAGTCCCCGAAGTGAAGGAGAATTCAAACAG
ACCTCGTCATTCTGGTGTGAAGCCTGGTTCGCTCACCGCCTATCATCTGCATTGGCTTACT
CAGGTGCTACCGACTCTGGCCCTGATGTCTGTAGTTTACAGGATGCCTTATTGTCTTC
TACACCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTAATAATGTCAGCTATGTGCCCC
ATCTCTCTTCATGCCCTCCCTCCCTTTCTTACCCTGCTGAGTGGCTGGAACCTGTTTTAAA
GTGTTTTATTCCCATTTCTTTGAGGGATCAGGAAGGAATCTGGGTATGCCATTGACTTCCC
TTCTAAGTAGACAGCAAAATGGCGGGGTCGAGGAATCTGCACTCAACTGCCACCTGGC
TGGCAGGGATCTTGAATAGGTATCTTGAGCTTGTTCTGGGCTCTTTCCTTGTGTACTGAC
GACCAGGGCCAGCTGTTCTAGAGCGGAATTAGAGGCTAGAGCGGCTGAAATGTTGTTTGG
TGATGACACTGGGGTCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG
CCACTGGGATCCCTCTGCCCTGTCTCTCTGAATACAAGCTGACTGACATTGACTGTGTCTGT
GGAAAATGGGAGCTCTTGTGTGGAGAGCATAGTAAATTTTACAGAGAACTGAAGCCAAAAG
GATTTAAAACCGTCTCTTAAAGAAAAGAAAACCTGGAGGCTGGGCGCAGTGGCTCACGCCCTG
TAATCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTTCGGGAGTTTCGGGATCAGCCTGACCA
ACATGGAGAAACCCCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCCTGTAGTC
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAACTCCAGCTCAAAAAAAAAAAAAAAAAA

FIGURE 44

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPKVLS CAYS GFSSPRVEW
KFDQGD TTRLVCYNNKITASYEDRVTF LPTGITFKSVTREDTGT YTCMVSEEGNSYGEVKV
KLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPEYTWFKDGI VMP TNP KSTRAF SNS
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGT P MTSNAVRMEAVERNVGVI VAAVLVT LI
LLGILVFGIWFAYS RGHFDR TKKGTS SKKVIYSQPSARSEGEFKQTSSFLV

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 238-255

N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 270-274

Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,
193-197, 203-207, 287-291

N-myristoylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

FIGURE 45

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGC**ATG**AGCGGCGGTTGGATGGCGCAGGTTGGA
GCGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTCGGCCTCGGACTAGGCCT
GGAGGCCGCCGCGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCCGCAGGCCCCAGCTCAG
GCTCGTGCCACCCACCAAGTTCCAGTGCCGCACCACTGGCTTATGCGTGCCCTCACCTGG
CGCTGCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGAGTGACAGGATTGAGCC
ATGTACCCAGAAAGGGCAATGCCACCGCCCGCTGGCCTCCCTGCCCTGCACCGCGCTCA
GTGACTGCTCTGGGGAACTGACAAGAACTGCGCAACTGCAGCCGCTGGCCTGCCTAGCA
GGCGAGCTCCGTTCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCA
CCCAGACTGTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGAAG
GGGATGCCACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCACCTCTCTCAGGAATGCC
ACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCCTCTGTGCGGAATGCCACATCCTC
CTCTGCCGAGACCACTGTGGAAGCCCAACTGCCCTATGGGGTTATGACGCTGCTGCGGTGC
TCAGTGCAAGCCTGGTACCGCCACCCCTCTCCTTTGTCTCGGCTCCGAGCCAGGAGCGC
CTCCGCCCACTGGGGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTGAGAACAAGAC
CTCGCTGCCCT**TG**AGGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACA
GGAGGAGAGCAGTGATGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTT
CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCC
TGGACACTCCCTATGGAGATCCGGGGAGCTAGGATGGGGAACTGCCACAGCCAGAACTGAG
GGGCTGGCCCCAGGCAGCTCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTG
CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

FIGURE 46

MSGGWMAQVGAWRTGALGLALLLLGLGLGLEAAASPLSTPTSQAAGPSSGSCPPTKQCR
TSGLCVPLTWRCRDLDCSDGSDEEECRIEPCTQKGQCPPPPGLPCPCTGVSDCSGGTDKKL
RNC SRLACLAGE LRCTLSDDCIPLTWRC DGHDCPDSSDELGCGTNEILPEGDATTMGPPVT
LESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAVLSASLVTATLL
LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

Signal sequence:

amino acids 1-30

Transmembrane domain:

amino acids 230-246

N-glycosylation site.

amino acids 126-130, 195-199, 213-217

Casein kinase II phosphorylation site.

amino acids 84-88, 140-144, 161-165, 218-222

N-myristoylation site.

amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218,
224-230, 230-236, 263-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 44-55

Leucine zipper pattern.

amino acids 17-39

FIGURE 47

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGG
GTTAGACTGGCGGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGA
CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGATATCATGGAATGAACCCGAGCAATG
GAGATGGATTCTTAGAGCAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTTG
GCCGTGATCCTGTGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTT
CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCG
GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCCGATTCTACTGCCAAGACGGATTCAAGCTGAAG
GGCGCTACAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCTAGGCTGGATCCCAAGTGA
TAATTCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCATA
ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTTCATGAAGGATTCAAGATCCGG
TACCCCGACCTACACAATATGGTTTCATTATGTGCGGATGATGGAACGTGGAATAATCTGCC
TACTCTGTCAAGGCTCGCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC
TCCAGACCTCCTTCCCGGTGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTT
GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCCGGTG
CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCTGCT
GCCACCCGCGCCTTGTGAGCGCTACAACCACGGAACGTGTGGTGGAGTTTTACTGCGATCCT
GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTCTCTC
TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCGACGCCATGAGACCTCCTGA
CCACGTGGAAGATTGTGGCGTTCACGGCAACCAGTGTGCTGCTGGTGTCTGCTCGTCATC
CTGGCCAGGATGTTCCAGACCAAGTTC AAGGCCACTTTCCCCCAGGGGGCCTCCCCGGAG
TTCAGCAGTGACCTGACTTTGTGGTGGTAGACGGCGTGCCCGCTCATGCTCCCGTCTTATG
ACGAAGCTGTGAGTGGCGGCTTGAGTGCCTTAGGCCCGGGTACATGGCCTCTGTGGGCCAG
GGCTGCCCTTACCCGTGGACGACCAGAGCCCCCAGCATACCCCGGCTCAGGGGACACGGA
CACAGGCCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA
GCTGTGATTACCTCCCAAGTGCC AAGAGGACACCCACCTGCTTCGGACAAACCTGACATA
ATTGCCAGCAGGCGAGGAGGTGGCATCCACAGCCAGGCATCCATCATGCCCACTGGGT
GTTGTTCCCTAAGAACTGATTGATTAAAAAATTTCCCAAAGTGTCTGAAGTGTCTCTTCAA
ATACATGTTGATCTGTGGAGTTGATTCTTTCTCTCTGTTTATAGACAAATGTAAACAA
AGCTCTGATCCTTAAAAATTGCTATGCTGATAGAGTGGTGGGCTGGAAGCTTGATCAAGTC
CTGTTTCTTCTTGACACAGACTGATTAAAAAATTAAGNAAAAAA

FIGURE 48

MYHGMNPSNGDGFLEQQQQQQPQSPQRLLAVILWFLALCFGPAQLTGGFDDLQVCADPGI
PENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQEDCRI
PQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLAS
SNGYVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPRCLALEAQVCPLP
PMVSHGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQT
WPSTHETLLTTWKIVAFTATSVLLVLLLVILARMFQTKFAHFPPRGPSPRSSSDPDFVVVD
GVPVMLPSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGD TDTGPGSESETCDS
VSGSSELLQSLYSPPRCQESTHPASDNPDIIASTAEVASTSPGIHHAHWVFLRN

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 325-344

N-glycosylation site.

amino acids 104-108, 134-138, 192-196

Casein kinase II phosphorylation site.

amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366,
364-368, 380-384, 467-471, 468-472

N-myristoylation site.

amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424,
478-484

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 394-405

FIGURE 49

CCCACGCGTCCGCTCCGCGCCCTCCCCCGCCTCCCGTGCGGTCCGTCCGGTGGCCTAGAGA
 TGCTGCTGCCGCGGTTGCAGTTGTGCGGCACGCCTCTGCCCGCCAGCCCGCTCCACCGCCGT
 AGCGCCCGAGTGTGGGGGGCGCACCCGAGTCGGGCC**ATG**AGGCCGGGAACCGCGCTACAGG
 CCGTGCTGCTGGCCGTGCTGCTGGTGGGGCTGCGGGCCGCGACGGGTGCGCTGCTGAGTGCC
 TCGGATTTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGGAGGGACACAGAGGCCCTTGTTA
 TAAAGTCATTTACTTCCATGATACTTCTCGAAGACTGAACCTTTGAGGAAGCCAAAGAAGCCT
 GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAAACTGATAGAA
 AAGTTCATTGAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCCGTGA
 GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGACTGATGGCAGCATAT
 CACAATTTAGGAACTGGTATGTGGATGAGCCGTCTTGC CGGCAGCGAGGTCTGCGTGGTTCATG
 TACCATCAGCCATCGGCACCCCGCTGGCATCGGAGGCCCTACATGTTCCAGTGGAATGATGA
 CCGGTGCAACATGAAGAACAAATTTCAATTTGCAAATATCTGATGAGAAACCAGCAGTTCCTT
 CTAGAGAAGCTGAAGGTGAGGAAACAGAGCTGACAAACCTGTACTTCCAGAAGAAACACAG
 GAAGAAGATGCCAAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACAT
 CCTAATCCCCAGCATTCCCCCTTCTCCTCCTCTTGTGGTCACCAAGTTGTATGTTGGGTTT
 GGATCTGTAGAAAAAGAAAAACGGGAGCAGCCAGACCTAGCACAAAGAAGCAACACACCATC
 TGGCCCTCTCCTCACCAGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAAACA
 AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTCAATTCGAGTGTGTT
 CGGGAGAAGCCACTCCCGATGACATGTCTTGTGACTATGACAAATGGCTGTGAACCCATCA
 GAAAGTGGGTTTGTGACTCTGGTGAAGCTGGAGAGTGGAATTTGTGACCAATGACATTTATGA
 GTTCTCCCCAGACCAATGGGGAGGAGTAAGGAGTCTGGATGGGTGGAAATGAAATATATG
 GTTAT**TAG**GACATATAAAAAAAGTAACTGACAACAATGGAAAAAGAAATGATAAGCAAAATC
 CTCTTATTTTCTATAAGGAAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGT
 GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCAAGTTTGGCTGTATCCTTTAT
 CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCAGGTCTGGCAGATAGTA
 GAGTCTCAATAAATGTCACTTGGTTGGTTGTATCTAACTTTAAGGGACAGAGCTTTACCTG
 GCAGTGATAAAGATGGGCTGTGGAGCTTGAAAAACCACTCTGTTTTCTTGCTCTATACAG
 CAGCACATATATCATACAGACAGAAAAATCCAGAATCTTTCAAAGCCACATATGGTAGCACAG
 GTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTTCAAAGAAATAAAATCAAATAAAGA
 GCAGGAAAAA

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 50

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGQPVCRRGTQRPCYKVIYFHDTSRRL
NFEEAKEACRRDGGQLVSI ESEDEQKLEKFIENLLPSDGDWFIGLRRREEKQSNSTACQDL
YAWTDGSI SQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY
SDEKPAVPSREAEGETELTTPVLPEETQEEDAKKTFKESREALNLAYILIPSIPLLLLLLV
VTTVVCWWVICRKRKREQPDSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRPDL
KNISFRVCSGEATPDDMSCDYDNMAVNPSES GFVTLVSVESGFVTNDIYEFSPDQMGRSKES
GWVENEIYGY

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 117-121, 312-316

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 296-300

Casein kinase II phosphorylation site.

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226,
299-303, 306-310, 323-327

N-myristoylation site.

amino acids 18-24, 37-43, 76-82, 146-152

FIGURE 51

GGGGTCTCCCTCAGGGCCGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTACGC
ATCCGAGGTTCCTCGCGACTTGGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGT
GTTTGCCCTCCTGCAGCCTCAACCCGAGGGCAGCGAGGGCTACACCATGATCACTGGTGT
GTTCAGCATGCGCTTGTGGACCCAGTGGGCGTCTGACCTCGTGGCGTACTGCGCTGCACC
AGCGGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGC
CTGCTGAAGTTGAAAATGGTGCAGGTCTGTTTCGACACGGGGCTCGGAGTCTCTCAAGCC
GCTCCCGCTGGAGGAGCAGGTAGAGTGGAAACCCCGACTATTAGAGGTCCCACCCCAAACCTC
AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT
CAATACCATGAGACCACCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCAT
GCAGCAAATGTTTGCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCTTTTC
TTTCACCAACCTTCAACCCACAGGAGGTCTTTATTGCTTCCACTAACATTTTTCGGAATCTG
GAGTCCACCCGTGTTTGGCTGGGCTTTTCCAGTGTGAGAAAGAGGCCATCATCAT
CCACACTGATGAAGCAGATTGAGAAGTCTGTATCCCAACTACCAAAGCTGCTGGAGCCTGA
GGCAGAGAACAGAGGCCGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTG
AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCTCT
GGACAACGTGGCTGCCGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTG
CACGGATGATCGAACAGAGAGCTGTGGACACATCCTTGACATACTGCCAAGGAAGACAGG
GAAAGTCTTCAGATGGCAGTAGGCCATTCTCCACATCTAGAGAGCAACCTGTGTAAAGC
CATGGACTCTGCCACTGCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG
TGACCTTCATACCGCTCTTAATGACCCTGGGGATTTTGACCACAAATGGCCACCGTTTGCT
GTTGACCTGACCATGGAACCTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA
TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCGCTGGACATGT
TCTTGAATGCCATGTGAGTTTATACCTTAAGCCAGAAAAATACCATGCACCTCTGCTCTCAA
ACTCAGGTGATGGAAGTTGGAATGAAGAGTAACTGATTATAAAAGCAGGATGTGTTGATT
TTAAATAAAGTGCCTTTATACAATG

FIGURE 52

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLMVQVVFRHGAR
SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQL
TKVGMQQMFALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKE
GPILIIHTDEADSEVLYPNYQSCWSLRQRTGRRRQTASLQPGISEDLLKKVKDRMGIDSDDKVD
FFILLDNVAAEQAHNLPSCPMLKRFRMIEQRAVDTSLYILPKEDRESLQMAVGPFPLHILES
NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLMLTLGIFDHKWPPFAVDLTMELYQHLESKEW
FVQLYYHGKEQVPRGCPDGLCPLDMFLNAMSYYTLSPKEYHALCSQTQVMEVGNNE

Signal sequence:

amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 218-222

Casein kinase II phosphorylation site.

amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site.

amino acids 280-288

N-myristoylation site.

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.

amino acids 216-220

Leucine zipper pattern.

amino acids 10-32

Histidine acid phosphatases phosphohistidine signature.

amino acids 50-65

FIGURE 53

CTCCTCTTAACATACTTGCAGCTAAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT
TAAATTTTCAGCTCATCACCTTCACCTGCCTTGGTCATGGCTCTGCTATTCTCCTTGATCCTT
GCCATTTGCACCAGACCTGGATTCTCTAGCGTCTCCATCTGGAGTGGCGTGGTGGGGGGCCT
CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG
ACGGCTGGGACATTAAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGC
GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCA
ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG
ATTGTTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCAGAGAGCTCTTTCTCCCCA
GTCCACAGAGGTGTGAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACCGCTGGAAGTGAA
GCACCAGAACAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGCCGCAAAAGGTGG
TGTGCCCGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC
TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCTGCTCAGGACGAGAAGCAACCTTCA
GGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCATGATGAAGACCGTGGGTGCG
AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG
GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACCTGGGGAGAAAAGGAGGA
CCAGGTGGTATGCAAGCAACTGGGCTGTGGGAAGTCCCTCTCTCCCTCCTTCAGAGACCGGA
AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTGTTGCTCAGGGGAGGAG
CAGTCCCTGGAGCAGTGCCAGCACAGATTTTGGGGGTTTACGACTGCACCCACCAGGAAGA
TGTGGCTGTCTGTCTCAGTGTTAGGTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAA
GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT
TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGCCCTGATTCTCAGGCCCTTCAGAGTTGG
ATCAGAACCTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACTACATCA
CCACCTTTCCATGTCTCCACATTGCACACAGCAGATTCACAGCCTCCATAATTGTGTGTAT
CAACTACTTAAATACATTCTCACACACACACACACACACACACACACACACACATA
CACCATTTGTCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA
TCTTAGAGGAACGGAATTTTAAGGATAAAATTTCTGAATTGGTTATGGGGTTTCTGAAAATTG
GCTCTATAATCTAAATTAGATATAAAATCTGGTAACCTTTATTTACAAATTAATAAGATAGCAC
TATGTGTTCAAA

FIGURE 54

MALLFSLILAICTRPGFLASPSGVRVLVGGHLRCEGRVEVEQKGQWGTVCDDGWDIKDVAVLC
RELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEVEYDCSHDEDAGASC
ENPESSEFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCTGWSLRAAKVVCRLGCGRAVL
TQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPGWKNTCNHDEDTWVECEDPFDLRLVG
GDNLCSGRLEVLHKGVWGSVCDDNWGEKEDQVVCKQLGCGKSLSPSFRDRKCYGPGVGRIWL
DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

Signal sequence:

amino acids 1-15

Casein kinase II phosphorylation site.

amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238,
267-271, 294-298, 316-320, 336-340

N-myristoylation site.

amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143,
180-186, 263-269, 286-292

Amidation site.

amino acids 196-200

Speract receptor repeated domain signature.

amino acids 29-67, 249-287

FIGURE 55

ACTGCACTCGGTTCTATCGATTGAATTCCTCCCGGGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCGGACGCGTGGGCGGACGCGTGGGCGGCTACCCAGGAAGAGTCTGCCGAAG
GTGAAGGCCATGGACTTTCATCACCTCCACAGCCATCTGCCCTGCTGTTTCGGCTGCTTGGG
CGTCTTCGGCTCTTCCGGTGCCTGCAGTGGGTGCGCGGGAAGGCCTACCTGCGGAATGCTG
TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAAGAATGTGCAAAAGTCTTCTATGCT
GCGGGTGCTAAACTGGTGCTCTGTGGCCGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA
ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACACAAGCCTTACTTGGTGACCTTCGACC
TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC
GACATACTTGTCACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGGA
TGTGGACAAGAGGGTTCATGGAGACAACTACTTTGGCCAGTTGCTCTAACGAAAGCACTCC
TGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTGCCATCAGCAGCATCCAGGGCAAG
ATGAGCATTCCTTTTTCGATCAGCATATGCAGCCTCCAAGCACGCAACCCAGGCTTTCTTTGA
CTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA
TCCACACCAACCTCTCTGTAATGCCATCACCGCGGATGGATCTAGGTATGGAGTTATGGAC
ACCACCACAGCCCAGGGCCGAAGCCCTGTGGAGGTGGCCCAGGATGTTCTTGCTGCTGTGGG
GAAGAAGAAGAAAGATGTGATCCTGGCTGACTTACTGCCTTCCTTGGCTGTTTATCTTCGAA
CTCTGGCTCCTGGGCTCTTCTTCAGCCTCATGGCCTCCAGGGCCAGAAAAGAGCGGAAATCC
AAGAACTCCTAGTACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGC
TTACTCTACAAGGGACAGTTGCATTTGTTGAGACTTTAATGGAGATTTGTCTCACAAGTGGG
AAAGACTGAAGAAACACATCTCGTGCAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA
AGCTTCTTCCAGGGTGAGGGGAAACACTTAAGGAATAAATATGGAGCTGGGGTTTAACT
AAAACTAGAAATAAACATCTCAACAGTAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAG
AGTCGACCTGCAGAAGCTTGGCCGCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTAC

FIGURE 56

MDFITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAAGA
KLVLCGRNGGAL EELIRELTASHATKVQTHKPYLVTFDLTDSGAIVAAAAEILQCFGYVDIL
VNNAGISYRGTIMD TTVDDVKRVMETNYFGPVALTKALLPSMIKRRQGHIVAIS SIQ GKMSI
PFRSAYAASKHATQAFFDCLRAEMEQYBIEVTVISPGYIHTNLSVNAITADGSRYGVM DTTT
AQGRSPVEVAQDVLAAVGKKKDVILADLLPSLAVYLR TLAPGLFFSLMASRARKERKSKNS

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 104-120, 278-292

N-glycosylation site.

amino acids 228-232

Glycosaminoglycan attachment site.

amino acids 47-51

Casein kinase II phosphorylation site.

amino acids 135-139, 139-143, 253-257

Tyrosine kinase phosphorylation site.

amino acids 145-153, 146-153

N-myristoylation site.

amino acids 44-50, 105-111, 238-244, 242-248, 291-297

Amidation site.

amino acids 265-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 6-17

FIGURE 57

CCCACGCGTCCGCTGGTGTAGATCGAGCAACCTCTAAAAGCAGTTTAGAGTGGTAAAAAA
AAAAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGGATGAAATTTCTTCTGGACATCCTC
CTGCTTCTCCCGTTACTGATCGTCTGCTCCCTAGAGTCCTTCGTGAAGCTTTTTATTCTTAA
GAGGAGAAAATCAGTACC GGCGAAATCGTGTGATTACAGGAGCTGGGCATGGAATTGGGA
GACTGACTGCCATGAATTTGCTAAACTTAAAAGCAAGCTGGTTCTCTGGGATATAAATAAG
CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACTGGGTGCCAAGGTTTATACCTTTGT
GGTAGACTGCAGCAACCGAGAAGATATTACAGCTCTGCAAAGAAGGTGAAGGCAGAAAATTG
GAGATGTTAGTATTTTAGTAAATAATGCTGGTGTAGTCTATACATCAGATTTGTTTGGCTACA
CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA
GGCATTTCTTCTGCAATGACGAAGAATAACCATGGCCATATTGTCACTGTGGCTTCGGCAG
CTGGACATGTCTCGGTCCCTTCTTACTGGCTTACTGTTCAAGCAAGTTTGTCTGTGTTGGA
TTTTCATAAACTTTGACAGATGAACTGGCTGCCTTACAAATAACTGGAGTCAAAACAACATG
TCTGTGTCTTAATTTCTGTAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA
CTCTGGAACCTGAGGAAGTGGTAAACAGGCTGATGCATGGGATTTCTGACTGAGCAGAAGATG
ATTTTATTCCATCTTCTATAGCTTTTTTAACAACATTGGAAAGGATCCTTCTGAGCGTTT
CCTGGCAGTTTTTAAACGAAAAATCAGTGTTAAGTTTGATGCAGTTATTGGATATAAAATGA
AAGCGCAATTAAGCACCTAGTTTTTCTGAAAACGTATTACCAGGTTTAGGTGATGTCATCTA
ATAGTGCCAGAATTTTAAATGTTTGAACCTCTGTTTTTCTAATTATCCCCATTTCTTCAATA
TCATTTTGGAGCTTTGGCAGTCTTCAATTTACTACCACTGTGTTCTTTAGCCAAAAGCTGATT
ACATATGATATAAACAGAGAAATACCTTTAGAGGTGACTTTAAGGAAAATGAAGAAAAGAA
CCAAAATGACTTTATTAAATAAATTTCCAAGATTATTTGTGGCTCACCTGAAGGCTTTGCAA
AATTGTACCATAACCGTTTATTAAACATATATTTTATTTTGATTGCACCTAAATTTTGT
ATAATTTGTGTTTCTTTTCTGTTCTACATAAAATCAGAAACCTCAAGCTCTCTAAATAAAA
TGAAGGACTATATCTAGTGGTATTTCAATGAATATCATGAACTCTCAATGGGTAGGTTTC
ATCTACCCATTGCCACTCTGTTTCTGAGAGATACCTCACATTCCAATGCCAAACATTTCT
GCACAGGAAGCTAGAGGTGATACAGTGTGCAAGTATAAAAGCATCACTGGGATTTAAG
GAGAATTTGAGAGAATGTACCCACAAATGGCAGCAATAATAAATGGATCACACTTAAAAAAA
AA
AA

FIGURE 58

MKFLLDILLLLPLLVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSK
LVLWDINKHGLEETAACKCKGLGAKVHTFVVDCSNREDIYSSAKKVKAETGDSILVNNAGVV
YTSDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYC
SSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVNRLMH
GILTEQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMQ

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34, 283-287

Casein kinase II phosphorylation site.

amino acids 52-56, 95-99, 198-202, 267-271

N-myristoylation site.

amino acids 43-49, 72-78, 122-128, 210-216

FIGURE 59

CCCACGCGTCCGCGGACGCGTGGGTCGACTAGTTCTAGATCGCGAGCGGCCGCCCGCGGCTC
AGGGAGGAGCACCGACTGCGCCGACCCCTGAGAGATGGTTGGTGCCATGTGGAAGGTGATTG
TTTCGCTGGTCCGTGTGATGCCTGGCCCCGTGTGATGGGCTGTTTCGCTCCCTATACAGAAGT
GTTTCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGC
TGGGAAGATCCAAAAAGGAAGAGAATTGAGTTTGGTCGGCCCTTTCCAGGACTGAACATGA
AGAGTTATGCCGGCTTCTCACCCGTGAATAAGACTTACAACAGCAACCCTTCTTCTGGTTT
TTCCAGCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGG
AGGTTTCATCCTGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTCAACAGTAACATGA
CCTTGGCGTGACAGAGACTTCCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCA
GTGGGCACAGGCTTCACTTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGT
AGCACGGGATTTTATACAGTGCCTAATTCAGTTTTCAGATATTTCCTGAATATAAAAAATA
ATGACTTTTATGTCACTGGGGAGTCTTATGACAGGAAATATGTGCCAGCCATTGCACACCTC
ATCCATTCCTCAACCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA
TGGATATTCTGATCCCGAATCAATTATAGGGGGCTATGCAGAATTCCTGTACCAAATTGGCT
TGTTGGATGAGAAGCAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATC
AGGAAGCAGAAGCTGGTTTGAGGCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAAC
AAGTGATCCTTCTTACTTCCAGAATGTTTACAGGATGTAGTAATTACTATAACTTTTTCGGGT
GCACGGAACCTGAGGATCAGCTTTACTATGTGAATTTTGTCACTCCAGAGGTGAGACAA
GCCATCCACGTGGGGAATCAGACTTTTAATGATGGAACATAGTTGAAAAGTACTTGCAGAGA
AGATACAGTACAGTCACTTAAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGA
TCTACAATGGCCAACTGGACATCATCGTGGCAGCTGCCCCGACAGAGCGCTCCTTGATGGGC
ATGGACTGGAAGGATCCAGGAATACAAGAAGGCAGAAAAAAAGTTTGGAGATCTTTAA
ATCTGACAGTGAAGTGGCTGGTTACATCCGCAAGCGGGTACTTCCATCAGGTAATTATT
GAGGTGGAGGACATATTTTCCCTATGACAGCCTCTGAGAGCTTTTGACATGATTAATCGA
TTCATTTATGGAAGGATGGGATCCTTATGTTGGATAAACTACCTCCCAAAAGAGAACAT
CAGAGGTTTTCATTGCTGAAAAGAAAATCGTAAAAACAGAAAATGTCTAGGAATAAAAAAA
TTATCTTTTTCATATCTGCAAGATTTTTTTCATCAATAAAAAATTACCTTGAAAACAGTGAGC
TTTTGTTTTTGGGGGGAGATGTTTACTACAAAATTAACATGAGTACATGAGTAAGAATTACA
TTATTTAACTTAAGGATGAAGGTATGGATGATGTGACACTGAGACAAGATGTATAAATGA
AATTTTAGGGTCTTGAATAGGAAGTTTAAATTTCTTCTAAGAGTAAGTGAAAAGTGCAGTTG
TAACAAACAAAGCTGTAACATCTTTTTCTGCCAATAACAGAAGTTTGGCATGCCGTGAAGGT
GTTTGGAAATATTATTGGATAAGAATAGTCAATTATCCCAATAAATGGATGAAGCTATAA
TAGTTTTTGGGAAAAGATTCTCAATGTATAAAGTCTTAGAACAAAAGAAATCTTTGAAATA
AAAATATTATATATAAAGTAAAAAATAA

FIGURE 60

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGDGQPLFLTPYIEAGKIQKGRELSL
VGPFPGGLNMKSYAGFLTIVNKTYNSNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEH
GPYVVTSNMTLRDRDPWTTLSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALIQF
FQIFPEYKNNDFYVTGESYAGKYVPAIAHLIHSNLPVREVKNLNGIAIGDGYSDPESIIGG
YAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGLTSDPSYFQNVTG
CSNYYNFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNTQTFNDGTIVEKYLRDITVQSVKPWLT
EIMNNYKVLIIYNGQLDIIVAAALTERSLMGMDWKGSGEYKKAEEKVVKIFKSDSEVAGYIRQ
AGDFHQVIIRGGGHILPYDQPLRAFDMINRFIYKGWDPPYVG

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 81-85, 132-136, 307-311, 346-350

Casein kinase II phosphorylation site.

amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352,
353-357, 424-428

Tyrosine kinase phosphorylation site.

amino acids 423-432

N-myristoylation site.

amino acids 22-28, 110-116, 156-162, 232-238

Serine carboxypeptidases, serine active site.

amino acids 200-208

Crystallins beta and gamma 'Greek key' motif signature.

amino acids 375-391

FIGURE 61

CGAGGGCTTTTCCGGCTCCGAATGGCACATGTGGGAATCCCAGTCTTGTGGCTACAAACAT
 TTTTCCCTTTCCTAACAAGTTCTAACAGCTGTTTCTAACAGCTAGTGATCAGGGGTTCTTCTT
 GCTGGAGAAGAAAGGGCTGAGGGCAGAGCAGGGCACTCTCACTCAGGGTGACCAGCTCCTTG
 CCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGGAGTGAGGTGATGGAAG
 TCTAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC
 CTGGGGGAGGGCCTGCCTTAAACAAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGAT
 AAGACGTGCCGGTAGGATAGGGAAGACTGGGTTTAGTCTTAATATCAAATTGACTGGCTGGG
 TGAACTTCAAAGCCTTTTAACTCTCTGGGAGATGAAAACGATGGCTTAAGGGGCCAGAAA
 TAGAGATGCTTTGTAATAAATAAATTTTAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA
 GACCAAAATAGATAAACAGGATTCCCTGAACATTCTTAAGAGGGAGAAAAGTATGTTAAAAATA
 GAAAAACCAAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGGATGGGGACCCCTGGGTC
 AGGCCACGCTCTTTGCTCCTCCCGAAATTATTTTTGGTCTGACCCTCTGCCTTGTGTTTT
 GCAGAAATCATGTGAGGGCCAAACGGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCT
 CCTCACCGCGCCCTCTCAGCATGGAAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGG
 ACAGCCGCTCTGTGGTCTGCTCTCAGTGGTCTGGGTGCTGCTGGCCCCCAGCAGCCGGC
 ATGCTCAGTTCAGCTTCCACTCTGAGAATCGTGACTGGACCTTCAACCACTTGACCGT
 CCACCAAGGGACGGGGGCCGTCTATGTGGGGGCCATCAACCGGCTATAAGCTGACAGGCA
 ACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCG
 CCCCCTCATCGTGCAGCCTGCAGCGAAGTGCTCACCCCTCACCAACAATGTCAACAAGCTGCT
 CATCATTGACTACTCTGAGAACCGCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGGTCTGCA
 AGCTGCTGCGGCTGGATGACCTCTTCATCCTGGTGGAGCCATCCCAACAAGAGGAGCACTAC
 CTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATGTGTGCTCTGAGGGGTGAGGA
 TGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCGACCCTGTCCA
 GCCGAAGCTGCCCCGAGACCCTGAGTCCTCAGCCATGCTCGACTATGAGTACACAGCGAT
 TTTGTCTCCTCTCTCATCAAGATCCCTTCAGACACCCCTGGCCCTGGTCTCCACTTTGACAT
 CTTTACATCTACGGCTTTGTCTAGTGGGGGCTTTGTCTACTTTCTCACTGTCCAGCCCGAGA
 CCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTCTTCTACACCTCACGCATCGTGGCG
 CTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCCTTCGGCTGCACCCGGGC
 CGGGGTGGAATACCCGCTCTCTGAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCC
 AGGCCTTCAATATCACAGCCAGGACGATGTACTCTTGGCCATCTTCTCCAAGGGCAGAAG
 CAGTATCACCAACCCGCCGATGACTCTGCCCTGTGTGCCTTCCCTATCCGGGCCATCAACT
 GCAGATCAAGGAGCGCCTGCAGTCTGTCTACAGGGCAGGGCAACCTGGAGCTCAACTGGC
 TGCTGGGAAGGACGTCAGTGCACGAAGCGCCTGTCCCATCGATGATAACTTCTGTGGA
 CTGGACATCAACCAGCCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTGTACACCAC
 CAGCAGGGACCGCATGACCTCTGTGGCCTTACAGCTTACACCGGCTACAGCGTGGTTTTGT
 TGGGAGCTAAGAGTGGCAAGCTGAAAAAGGTAAAGAGTCTATGAGTTCAGATGCTCCAATGCC
 ATTCACTCCTCAGCAAAAGAGTCCCTCTTGAAGGTAGCTATTGGTGAGATTTAACTATAG
 GCAACTTTATTTTCTTGGGGAACAAAGGTGAATGGGGAGGTAAAGAGGGGTAAATTTTGTG
 ACTTAGCTTCTAGCTACTTCTCCAGCCATCAGTCATTGGGTATGTAAGGAATGCAAGCGTA
 TTTCAATATTTCCCAAACCTTAAAGAAAACTTTAAGAAGGTACATCTGCAAAAGCAA

FIGURE 62

MGTLGQASLFAPPGNYFWS DH SALCF AESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWP
RALEVDSRSVLLSVVWVLLAPPAAGMPQFSTFHSEN RDWTFNHLTVHQGTGAVYVGAINRV
YKLTGNLTIQVAHKTGFEE DNKSRY PPLIVQPCSEVLT LTN NVNKLLI IDYSENRL LACGSL
YQGVCKLLRLDDL FILVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVDGKQDY
FPTLSSRKLPRDP ESSAMLDYELHSD FVSSLIKIPSDTLALVSHFDIFYIYG FASGGFVYFL
TVQPETPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP
GDSLAQAFNITSQDDVLF AIFSKGQKQYHHPDDSALCAFP IRAINLQIKERLQSCYQGEGN
LELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNG
YSVVFVGTKSGK LKKVRVYEFRC SNAIHLLSKESLLEGSYWWRFN YRQLYFLGEQR

Signal sequence:

amino acids 1-32

Transmembrane domain:

amino acids 71-87

N-glycosylation site.

amino acids 130-134, 145-149, 217-221, 381-385

Casein kinase II phosphorylation site.

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387,
384-388, 471-475, 481-485, 530-534

N-myristoylation site.

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

FIGURE 63

AGGCTCCCAGCGCGGCTGAGTGGGACTGGAGTGGGAACCCGGTCCCCGCGCTTAGAGAACACGCGCATGACCA
 CGTGGAGCCTCCGCGGAGGCGCGGCCGACGCTGGGACTCCTGTGCTGGTGGTCTTGGGCTTCTGGTGTCTCC
 GCAGGCTGGACTGGAGCACCCTGGTCCCTCTGGGCTCGCCATCGACAGCTGGGGCTGCAGGCCAAGGGCTGGA
 ACTTCATGCTGGAGGATTCACCTTCTGGATCTTGGGGCTCCATCCACTATTTCCGTGTGGCCAGGAGTACT
 GGAGGACCGCTGCTGAAGATGAAGCCCTGGGCTTGAACACCTCACACCTATGTTCCGTGGAACTGCAATG
 AGCCAGAAAGAGGCAAAATTTGACTTCTCTGGGAACTGGACCTGGAGGCTTCTGTCTGATGGCCGACAGATCG
 GGCTGTGGGTGATTCTGCGTCCAGGCCCTACATCTGCAGTGAGATGGAACCTGGGGGCTTGGCCAGCTGGCTAC
 TCCAAGACCTGGCATGGAGCTGAGGACAACTTACAAGGGCTTCCACGAAGCAGTGGACCTTTATTTTGACACC
 TGATGTCCAGGGTGGTCCACTCAGTACAAGCGTGGGGGACCTATCATTTCCGCTCAGGTGGAGTGAATATG
 GTTCTCTATAAATAAGACCCCGCATACATGCCCTACGTCAAGAGGCACTGGAGGACCGTGGCATTTGTGGAACCTGC
 TCTGACTTCAAGACAACAGGATGGGCTGAGCAAGGGGATGTCCAGGGATCTTGGCCACCATCAACTTTCAGT
 CAACACACGAGCTGCAGCTACTGACCACCTTTCTCTCAACGTCAGGGGACTCAGCCCCAAGATGGTGTGAGT
 ACTGGACGGGTGGTTTGACTCTGGGGGAGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTGGAAAACCGTGT
 CTGCCATTGTGGACGCCGGCTCCTCCATCAACCTCTACATGTTCCACGGAGGCCAACCTTGGCTTCTGTAATG
 GAGCCATGCACCTTCCATGACTACAAGTCAGATGTCACGAGTATGACTATGATGCTGTGCTGACAGAAGCCGGCG
 ATTACACGGCCAGTACATGAAGCTTCGAGACTTCTTCGGCTCCATCTCAGGCATCCCTCTCCCTCCCCACCTG
 ACCTTCTTCCCAAGATGCGGTATGAGCCCTTAAACGCGAGTCTTGTACTGTCTCTGTGGGACGCCCTCAAGTACC
 TGGGGAGCCAAATCAAGCTGTAAGAGCCCATCAACATGGAGAACCTGCCAGTCAATGGGGGAAATGGACAGTCTCT
 TCGGGTACATTCTCTATGAGACCGAGCATCACCTGCTGGCATCCTCAGTGGCCACGTGTCATGATGGGGGCGAG
 TGTGTTGTAACACAGATCTCTATGAGGATTTCTGGACTACAAGACAACAGAGATGTGCTGCCCTTGATCAGAGGTT
 ACACGCTGTGAGGATTCCTAGGGAATCGTGGCGGAGTCAACTATGGGGAGAAATTATGATGACAGCGCAAAG
 GCTTAATTGGAAGTCTCTATCTGATGATTTCAACCCCTGAAAAACTTCAGAATCTATAGCCTGGATATGAAGAAGA
 GCTTCTTCAAGAGTTCGCGCTGGACAAATGSGNNTTCCCTCCAGAAACCCCACTATCCTGCTTCTCTCTTGG
 GTAGCTTGTCCATCAGCTCCAGCCTTGTGACACCTTTCTGAAGCTTGGAGGCTGGGAGAGGGGGTGTATTCA
 TCAATGGCCAGAACCTTGACGCTTACTGGAACATGGACCCAGAAAGCGCTTACCTCCAGGTCCCTGGTTGA
 GCAGCGGAATCAACAGGTCATCGTTTTTGGAGGACGATGGCGGGCTGCAATTACAGTTACGGAAGACCCCTCC
 ACCTGGGCAGGAACCGATGATTAACTGAGCGGCTGGCACCCCTCTCTGTGGTGGCCAGTGGGAGACTGCGCGCTC
 CTTTGAACCTCAAGCTTGTGGCTGTGCCCCACCCCTCACTGCAAAAGCATCTTAACTTAAGTAGCAACCTCAGGG
 ACTGGGGCTACAGTCTGCCCTGTCTCAGCTCAAAACCTTAAGCCTGCAGGGGAAAGGTGGGATGGCTCTGGGCC
 TGGCTTTGTTGATGATGGCTTCTTACAGCCCTGCTCTTGTGCCAGGCTGTGGGCTGTCTCTAGGGTGGGAGC
 AGCTAATCAGATCGCCAGCCTTTGGCCCTCAGAAAAAGTGTGAAACGTGCCCTTGCACCGGACGCTCAGAGCCC
 TGCAGCATCTGCTGGACTCAGGCGTGTCTTTGTGCTGTTCTGGGAGGCTTGGCCACATCCCTCATGGCCCAT
 TTATCTCCCGAAATCTGGTGTGTACCAAGTGTAGAGGTGGGGGAGGGGTGTCTACCTGAGCTGACTTGTGTT
 CTTCTTCAACAACCTTCTGAGCCTTCTTTGGGATTCTGGAAGGAACCTCGCGTGAGAAACATGTGACTTCCCCCT
 TCCCTTCCCATCTCGCTCTTCCACAGGGTGACAGGCTGGGCTGGAGAAACAGAAATCCTCACCTGCGCTCTCC
 CAAGTTAGCAGGTGTCTGTTGTTCAGTGAAGGAGCATGTGAGTCTGGCAGAGGCAATGGCCCATGTCTGCA
 CATCCAGGGAGGAGGACAGAAAGGCCAGCTCACATGTGAGTCTGGCAGAGGCAATGGCCCATGTCTGCAATCC
 AGGAGAGGAGGACAGAAAGGCCAGCTCACATGTGAGTCTGGCAGAGGCAATGGCCCATGTCTGCAATCCAGGAGG
 GGAGGACAGAAAGGCCAGCTCACATGTGAGTCTGGCAGAGGCAATGGCCCATGTCTGCAATCCAGGAGGAGG
 ACAGAAAGGCCAGCTCAGTGGCCCCCGCTCCCCACCCCCACGCGGCAACAGCAGGCGAGAGCAGCCCTCTTCT
 GAAGTGTGTCCAGTCCGATTTGAGCCTTGTCTGGGGCCAGCCCAACACCTGGCTTGGGCTCAGTGTCTGTA
 GTTGCAGTAAAGCTATAACCTTGAATCAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 64

MTTWSLRRRPPARTLGLLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFW
IFGGSIIHYFRVPREYWRDRLLKMKACGLNTLTITYVPWNLHEPERGKFDPSGNLDLEAFVLMA
AELGLWVILRPGPIYICSEMDLGGLPSWLLQDPMRLRTTYKGFTAEVDLYFDHLMRSRVVPLQ
YKRGGPILIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLLTSDNKGLSKGIQGVLAT
INLQSTHELQLLTTFLFNVQGTQPKMVMWEYWTGWFDSWGPHNILDSSSEVLKTVSAIVDAGS
SINLYMFHGGTNPFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDYTAHYMKLRDFFGSISGIP
LPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLFPVNGNGQSFGYILYE
TSITSSGILSGHVHDRGQVFVNTVSIQFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGEN
IDDQRKGLIGNLYLNDSPKKNFRIYSLDMKKSFFQRFGLDKWXSLEPETPLPAFFLGSLISIS
STPCDTFLKLEGWKEGVVFINQNLGRYWNIGPQKTLYLPGPWLSSGINQVIVFEETMAGPA
LQFTETPHLGRNQYIK

Signal sequence:

amino acids 1-27

Casein kinase II phosphorylation site.

amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

N-myristoylation site.

amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315,
320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

FIGURE 65

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCAGGACC
CTGGTGAGGGTTCTCTACTTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGG
GGAGCAAAGCCGGGCTCGGCCCGAGGCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC
CGACACGTGACGGTCTGTCTCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCCAATGGGTCCC
AAGAAGCTGTCTGCTTCGTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCA
GGCAGACACTCGGTTCGTTTCGTAGTGGATAGGGGTTCATGACCGGTTTCTCCTAGACGGGGCCC
CGTTCCGCTATGTGTCTGGCAGCCTGCACCTACTTTCGGGTACCGCGGGTGCTTTGGGCGGAC
CGGCTTTTGAAGATGCGATGGAGCGGCTCAACGCCATACAGTTTATGTGCGCTGGAACCTA
CCACGAGCCACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGA
ATGAGGCAGCTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAG
TGGGAGATGGGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCTC
AGATCCAGACTTCTTGGCGCAGTGGACTCCTGGTTCAAGGTCTTGTGCTCCCAAGATATATC
CATGGCTTTATCACAATGGGGGCAACATCATTAGCATTGAGGTGGAGAAATGAATATGGTAGC
TACAGAGCCTGTGACTTCAGCTACATGAGGCACTTGGCTGGGCTCTTCCGTGCACCTGTAGG
AGAAAAGATCTTGTCTTCCACACAGATGGGCCCTGAAGGACTCAAGTGTGGCTTCCCTCCGGG
GACTCTATACCACCTGTAGATTTTGGCCAGCTGACAACATGACCAAAATCTTTTACCCTGGCTT
CGGAAGTATGAACCCCATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTA
CTGGGGCCAGAATCACTCCACACGGTCTGTGTGCTGAGCTGTAACCAAAGGACTAGAGAATATGC
TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG
AATGGTGCCGATAAAGAGGGACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACC
TATATCTGAAGCAGGGGACCCACACCTAAGCTTTTGGCTCTTCGAGATGTATCAGCAAGT
TCCAGGAAGTTCCTTTGGGACCTTTACCTCCCCGAGCCCAAGATGATGCTTGGACCTGTG
ACTCTGCACCTGGTTGGGCATTACTGGCTTTCCTAGACTTGTCTTGGCCCCGTGGGGCCAT
TCATTCAATCTTGCCAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC
GAACTTATATGACCCATACCATTTTGTAGCCAACACCATTCTGGGTGCCAAATAATGGAGTC
CATGACCGTGCTATGTGATGGTGGATGGGGTGTTCAGGGGTGTGTGGAGCGAAATATGAG
AGACAAACTATTTTTCAGCGGGGAACTGGGGTCAAACCTGGATATCTTGGTGGAGAACATGG
GGAGGCTCAGCTTTGGGTCTAAACAGCAGTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTG
GGGCAACCAATCCTTACCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG
GTGGTTTCCCTCCAGTTGCCAAAATGGCCATATCCTCAAGCTCCTTCTGGCCCCACATTCT
ACTCAAAACATTTTCAATTTTAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG
ACCAAGGGCCAAAGTCTGGATCAATGGGTTTAACTTGGGCCGCTTGGACAAAAGCAGGGGCC
ACAACAGACCTCTACGTGCCAAGATTCTGTGTTTCTTAGGGGAGCCCTCAACAAATTA
CATTGCTGGAACGTAGAAGTGTACCTCTCCAGCCCCAAGTCCAATTTTGGATAAGCCTATC
CTCAATAGCAGTAGTACTTTGACAGGACACATATCAATTCCTTTTACGCTGATACACTGAG
TGCCTCTGAAACCAATGGAGTTAAGTGGGCACTGAAAGGTAGGCCGGGCATGGTGGCTCATGC
CTGTAATCCAGCACTTTGGGAGGCTGAGACGGGTGGAATTAACCTGAGGTGAGACTTCAAGA
CCAGCCTGGCCAACATGGTGAACCCCGTCTCCACTAAAAATACAAAATTAGCCGGGCGTG
ATGGTGGGCACCTCTAATCCAGCTACTTGGGAGGCTGAGGGCAGGAGAAATGCTGAATCC
AGGAGCAGAGGTTCCAGTGAGTGGAGGTTGTACACTGCACCTCCAGCTGGCTGACAGTGA
GACACTCCATCTCAAAAAAAAAA

FIGURE 66

MAPKKLSCLRSLLLPLSLTLLLPQADTRSFVVDGRHDFLLDGAPFRYVSGSLHYFRVPRVL
WADRLCLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNGSRDLIAFLNEAALANLLVILRPGPYI
CAEWEMGGPLSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENE
YGSYRACDFSVMRHLAFLFRALLGEEKILLFTTDDGPEGLKCGSLRGLYTTVDGFPADNMTKIF
TLRLRYEPHGPLVNSEYGTGWLQDYWGQNHSTRSVSAVTKGLENMLKLGASVNMYPFHGGTGF
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPSPKMML
GPVTLHLVGHLLAFLDLCLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHTIFEPTPFWVFN
NGVHDRAYVMVDGVFQGVVERNMMDKLFLTGKLGSKLDILVENMGRLSFGSNSDDFKGLLKP
PILGQTILTQWMMFPLKIDNLVKKWFPLQLPKWPYPQAPSGPTFYSKTFPILGSVGDFTFLYL
PGWTKGQVWINGFNLGRYWTQGPQQTLYVPRFLLFPFGALNKITLLEEDVPLQPQVQFLD
KPILNSTSLHRTINSLSADTLSASEPMELSGH

Signal sequence:

amino acids 1-27

N-glycosylation site.

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 4-8

Casein kinase II phosphorylation site.

amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554,
603-607, 644-648

Tyrosine kinase phosphorylation site.

amino acids 191-198

N-myristoylation site.

amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233,
231-237, 274-280, 296-300, 307-313, 447-453, 484-490

FIGURE 67

GCTTTGAACACGCTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC
ACCCACAATATGGCCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT
TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTCAGGATACCTTTGAAGGAATATT
CTTTGAAAAAGTCAGAGAAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAACGATTTT
GCGTTCCTTCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTTGGTGTGTT
CTTGTCAGAAGTTAGTGAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG
AAAACTCAGGCAGCACATTTACGCAACGCCAGGACAAGCAGGAGTTGCATCTGTTTCATG
CTGTGCGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAAC
AATTCAGAAGCTAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC
TCTGCCACTGCCCTGCAAAAGTTGAACAGACTGCTTTTAGCTTTCTTCGCGATCACTTGAGA
TGCCTTCACGTGAAGTTCACATGATGTGGCTGAAATTCCTGCCTGGGTGTATTTGCTCAAAAA
CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACCTCGAAAAAATAAGATGATAGGACTTG
AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTTGACCAAAGTT
CCCTCCAACATTACAGATGTGGCTCCACATCTTACAAAGTTAGTCATTACATAATGACGGCAC
TAAACTCTTGGTACTGAACAGCCTTAAGAAAAATGATGAATGTGCTGAGCTGGAACTCCAGA
ACTGTGAGCTAGAGAGAATCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACTGGAT
TTAAAGTCCAATAACATTGCGACAATTGAGGAAATCATCAGTTTCCAGCATTTAAAACGACT
GACTTGTTTAAAATTATGGCATAACAAAATTGTTACTATTCTCTCCCTCTATTACCCATGTCA
AAAACCTGGAGTCATTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT
AGTTTACAGAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA
AATAGGATTGCTTCAGAACCTGCAGCATTTGCATATCACTGGGAACAAAGTGGACATTTCTGC
CAAAACAATTTGTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAAGCTGCATCACC
TCACTCCCAGAGAAAGTTGGTCAGCTCTCCAGCTCACTCAGCTGGAGCTGAAGGGGAAC
CTTGAGACCGCCTGCCAGCCAGCTGGGCCAGTGTGCGATGCTCAAGAAAAGCGGGCTTGTG
TGGAAGATCACCTTTTTGATACCTGCCACTCGAAGTCAAGAGGCATTGAATCAAGACATA
AATATTCCTTTTGCAAAATGGGATTTTAAACTAAGATAATATATGCACAGTGATGTGCGGAAC
AACTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG
ATACATCTTTTAAAAATAAACAGAGAGGATGCATAGAAGGCTGATAGAAGACATAAAGTGAAT
GTTCAATGTTTGTAGGGTTTAAAGTCATTCAATTTCCAAATCATTTTTTTTTTTCTTTTGGGG
AAAGGGAAGGAAAAATATAATCACTAATCTTGGTTCTTTTAAATGTTTGTAACTGGAT
GCTGCCGCTACTGAATGTTTACAAATTGCTTGCCTGCTAAAGTAAATGATTAAATTGACATT
TTCTTACTAAAAAAAAAAAAAAAAA

FIGURE 68

MAYMLKKLLISYISIIICVYGFCLYTLFWLFRIPLKEYSFKEKREESSFSIDIPDVKNDFAFLLHMVDQYDQLYSKRFGVFLSEVSENKLEISLNHEWTFEKLROHISRNAQDKQELHLFMLSGLVPDAVFDLTDLVLKLELIPEAKIPAKISQMTNLQELHLCHCPAKVEQTAFSFLRDHLRCLHVKFTDVAEIPAWVYLLKNLRELYLIGNLSENKMMIGLESRLRELHLKILHVKSNLTKVPSNITDVAPHLTKLVIHNDGTKLLVLNSLKKMMNVAELELQNCLELIPHAIFSLSNLQELDLKSNNIRTIEEIIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLESLYFSNNKLES LPVAVFSLQKLRCLDVSYNNISIMPIEIGLLQNLQHLHITGNKVDILPKQLFKCIKLRTLNLGQNCITSLPEKVGQLSQLTQLELKGNCLELDRPLPAQLGQCRMLKKSGLVVEDHLFDFTLPLEVKEALNQDINIPFANGI

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 241-245, 248-252, 383-387

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 326-330

Casein kinase II phosphorylation site.

amino acids 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

Tyrosine kinase phosphorylation site.

amino acids 349-355, 375-381

N-myristoylation site.

amino acids 78-84, 124-130, 212-218, 392-398

FIGURE 69

CCCACGCGTCGGGCCCTTCTCTGGACTTTGCATTTCCATTCTTTCTATGACAACTGACTTTTTTTATTCTT
 TTTTTTCCATCTCTGGGGCAGCTTGGGATCTTAGGCCGCCCTGGGAAGACATTTGTGTTTACACACATAAGGAT
 CTGTGTGTTGGGGTTCTTCTTCTCCCTGACATTGGCATGCTTAGTGGTGTGTGGGGAGGGAGACCACGTGG
 GCTCAGTGTCTGTCTGCACTTATCTGCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGTATATGCTGTCT
 ATCGCTGGTGTATCTGTGGGCCCTTGCTCTGCTGATAGTTGTGCTGTCTGTCTTTATCTTCAAAATACACAAC
 GCGCTAAAGCTGCAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACACCCAGACAGGTGTGGTGGGCCAAG
 AACAGCCAGGCCAAAACCATTTGCCACGGAGTCTTGTCTGCCCTGCAGTGTCTGGAAGGATATAGAATGTGTGCC
 AGTTTGTGATTCCTCGCCACCTTGTGTGTGCGACATAAATGAGGGCCCTCTGAGTTAGGAAAGGCTCCCTTCTCAA
 GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTGTGATGTGCGAGCACAGAAAGAAAGGCACAG
 CTCCTCATCAGTTTATGTAAGAAATACTCAGTGCCTGCTGGGAACACAGCTGTGGAAGATCCCTACAGAGAGCTTC
 CACTGGGGCAACCCCTTCAGGAAGGAGTTGGGGAGAGAGAACCTCACTGTGGGGAAATGCTGATAAAACGATCA
 CACAGCTGCTCTATTCTCAACAATCTACCCCTTGCGTGGCTGGAACTGACGCTTTCCCTGGAGGTGTCCAGAAA
 GCTGATGTAACACAGAGCCTATAAAAGCTGTCGGTCCCTTAAGGCTGCCAGCGCTTGCCAAAATGGAGCTTGTA
 AGAAGGCTCATGCCATTGACCCCTCTTAATTTCTCTCCTGTTTGGCGGAGCTGACAAATGGCGGAGGCTGAAGGCAAT
 GCAAGCTGCACAGTCACTTAGGGGGTGCCAAATATGGCAGAGACCCACAAAGCCATGATCCTGCAACTCAATCCC
 AGTGAGAACTGCACCTGGCAATAGAAAGACAGAAAAACAAAGCATCAGAATTATCTTTTCTTATGTCCAGCTT
 GATCCAGATGGAAGCTGTGAAAGTGAACACATTAAAGTCTTTGACGGAACCTCAGCAATGGGCCCTCTGCTAGGG
 CAAGTCTGCAGTAAAAACGACTATGTTCTGTATTTGAATCATCATCCAGTACATTGACGTTTCAAATAGTTACT
 GACTCAGCAAGAATTCAAAGAACTGCTTTGTCTTTACTACTTCTCTCTCCAACTCTCTATTCCAACTGT
 GCGGTGTACCTGGATACCTTGGAGGATCTTCAACAGCCCCAATTACCCAAAGCCGCTCCTGAGCTGGCTTAT
 TGTGTGTGGCACAATAAGTGGAGAAAGATTACAAGATAAACTAAACTCAAAGAGATTTCTCTAGAAATAGAC
 AAACAGTGCAAATTGATTTCTTGCCCATCTATGATGGCCCCCTCAACCAACTCTGGCCTGATTTGGACAACTGTG
 GGCCGTGTGACTCCCACTTCGAATTCGATCAAATCTCTGACTGTGGTGTGTACAGATTATGCCAATTTCA
 TACCGGGGATTTTCTGCTTCTACACCTCAATTTATGACAGAAACATCAACCATACATCTTTAACTTGTCTCTTCT
 GACAGATGAGAGTTATTATAGCAAAATCTCACTAGAGGCTTTAACTCTAAATGGGAATAAATGCAAGTGAACAAA
 GACCCAACTTGACAGACAAAATATTCAAAATGTGTGGAATTTTCTGTCCCTCTTAATGGATGTGGTACAATCAGA
 AAGGTAGAAGATCAGTCAATTACTTACACCAATAATACCTTTCTGATCTCCAACTCTGGAAGTGAATCACC
 CGTCAGAAACAACTCCAGATTATTGTGAAGTGTGAAGTGGGACATAATTTCTACAGTGGAGATAATATACATAACA
 GAAGATGATGTAAATCAAAGTCAAATGCACTGGGCAAAATACACACGATGCTCTTTTGAATCCAAATTC
 TTTGAAAAGCATATACTTGAATCACCATAATATGTGGATTGAAACAACTCTTTTGTCTCAAGTTAGCTGAC
 ACCTCAGATCAAATTTGTGTGTGTTTCTGATACCTGTAGAGCCTCTCCCACTCTGACTTTGCATCTGCAACC
 TAGCACTTAATCAAGATGGAGTGTGATCGAGATGAACTTGAAGGTGATCCCTTATTTGGACATATGGGAGA
 TTCCAGTTTATATGCTTTTAAATCTTGTGAGAGATGAGCTCTGTGTATCTGCAGTGTAAAGTTTGTATATGTGAT
 AGCAGTGACCACAGCTCTCGCTGCAATCAAGGTTGTGTCTCAGAAAGCAACGAGACATTTCTTCAATATAAATGG
 AAAACAGATTCCATCATAGGACCACTTCTGTCTGAAAAGGGATCGAAGTGCAGTGGCAATTACAGGATTTACAGAT
 GAAACACATGCGGGAAGAACTCCAAACCAAGCCTTTCAACAGTGTGCATCTGTTTTCTTCTCATGTGTTCTAGCTCTG
 AATGTGGTGACTGTAGCGACATACAGTGGAGCATTTGTAAATCAACGGGAGCACTACAAATACCAAGAGCTG
 CAGAACATTAACCTAAGCTCCAACTTAAGTGAGACATGTTTTCTCAGGATGCGAAAGGAATGCTACCTCGT
 GGCTACACATTAATGAAATAATGAGGAAGGCCGTGAAGTGAACACAGGCCCTGATGTAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 70

MELVRRRLMPLTLLILSCLAELTMAEAEAGNASCTVSLGGANMAETHKAMILQLNPSENCTWTIERPENKSIRIIFSYVQLDDPDGSCSEENIKVFDGTSSNGPLLQVCSKNDYVPVFESSSSSTLT
FQIVTDSARIQRTVFVFFYFFSPNISIPNCGGYLDLTLEGSFTSPNYPKHPPELAYCVWHIQV
EKDYKIKLNFKEIFLEIDKQCKFDFLAIYDGPSTNSGLIGQVCGRVTPTFESSNSLTVVLS
TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVIIISKSYLEAFNSNGNNLQDKDPTCRP
KLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGHNST
VEIIYITEDDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPN
LVVFLDTCRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKPLRSMSSVYL
QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDISIIGPIRLKRDRSASGNSGFGQHE
THA BETPNQFPNSVHLFSFMYLALNVVTVATITVRHFVNQRADYKYQKLQNY

Signal sequence:

amino acids 1-24

Transmembrane domain:

amino acids 571-586

N-glycosylation site.

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374,
394-398, 419-423

Casein kinase II phosphorylation site.

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383,
408-412, 463-467, 520-524, 556-560

Tyrosine kinase phosphorylation site.

amino acids 172-180, 407-415, 407-416, 519-528

N-myristoylation site.

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 71

GACGGAAGAACAGCGCTCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCG
GGACATGCGGCCCCAGGAGCTCCCGAGGCTCGCGTTCCCGTTGCTGCTGTTGCTGTTGCTGC
TGCTGCCCGCGCCCGCTGCCCTGCCACAGCGCCACGCGCTTCGACCCACCTCGGGAGTCC
CTGGACGCCCGCAGCTGCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG
GGGAGTGTGTTTCCGTGCCAGCTTCGGTAGCGAGTGGTTCTGGTGGTATTTGGCAAAAGGAAA
AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCTCTCTAGTTTCAAATATGAAGAT
TTTGGACCACTATTACAGCAAAATTTTTTAATGCCAACCAAGTGGGCAGATATTTTTCAGGC
CTCTGGTGCCAAATACATTGTCTTAACTTCCAACATCATGAAGGCTTTACCTTGTGGGGGT
CAGAATATTCTGGAACTGGAATGCCATAGATGAGGGGCCCAAGAGGGACATTGTCAAGGAA
CTTGAGGTAGCCATTAGGAACAGAAGTACCTCGGTTTGGAGTGTACTATTCCTTTTTTGA
ATGGTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTTCATTCCATAAGCGGCAATTTCCAG
TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACTATCAGCCTGAGGTTCTGTGG
TCGGATGGTGACGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT
ATATAATGAAAGCCAGTTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA
TCTGTAAGCATGTTGGCTTCTATACCTGCAGTGTATCGTTATAACCCAGGCATCTTTTGCCA
CATAAATGGAAAACTGCATGACAATAGACAACTGTCTGGGGCTATAGGAGGGAAGCTGG
AATCTCTGACTATCTTACAATTGAAGAAATGGTGAAGCAACTGTAGAGACAGTTTCATGTG
GAGGAAATCTTTTGATGAATATTTGGGCCACACTAGATGGCACCAATTTCTGTAGTTTGTGAG
GAGCGACTGAGGCAAGTGGGGTCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA
TACCTGGCGATCCAGAAATGACACTGTACCCCCAGATGTGTGGTACACATCCAAGCCTAAAG
AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCCCATCAGGACAGCTGTCTCTTGGCCAT
CCCAAGCTATTTCTGGGGGCAACAGAGGTGAAACTACTGGGCCATGGACAGCCACTTAACTG
GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTCATCAGATGC
CGTGTAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTG
ATGCTGCAAGTTATGTCTAAGGCTAGGAACATCAGGTGTCTATAATTGTAGCACATGGAGA
AAGCAATGTAACTGGATAAGAAAAATATTTTGGCAGTTGAGCCCTTCCCTTTTTCCCACTA
AATTTTTCTTAAATTACCATGTAAACATTTTAACTCTCCAGTGCACTTGGCCATTAAAGTC
TCTTCACATTGATTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTTCACATTATAGTAG
CAAGGAATTTGGTGGTATTATGGACCGAACTGAAAAATTTTATGTTGAAGCCATATCCCCCATG
ATTATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCATTGTCTAGTCAAT
TTTTTTTTTGTGCCAACATCATAGATGTATTTACAAAAATCCTAGATGGCATAGCCTACTACA
CACCTAATGTGTATGGTATAGACTGTGCTCCTAGGCTACAGACATATACAGCATGGTTACTG
AATACTGTAGGCAATAGTAACAGTGGTATTGTTATATCGAAACATATGGAACATAGAGAAG
GTACAGTAAAAATACTGTAAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAG
CTTAGCAGGCTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTGAAGGCTTAGGACATTA
TGAAACACTGCCAGACGTTATAAAATACTGTATGCTTAGGCTACACTACATTATAAAAAAA
GTTTTTCTTTCTTCAATTATAAAATTAACATAAGGTGACTGTAACTTTACAAACGTTTTTAATT
TTTAAACCTTTTTGGTCTTTTGTAAACACTTAGCTTAAACATAAACTCATTTGTGCAA
ATGTAA

FIGURE 72

MRPQELPRLAFPLLLLLLLLLPPPPCPAHSATRFDPWESLDARQLPAWFDQAKFGIFIHWG
VFSVPSFGSEFWWWYQKEKIPKYVEFMKDNYPFSFKYEDFGPLFTAKFFNANQWADIFQAS
GAKYIVLTSKHHEGFTLWGSSEYSWNWNAIDEGPKRDIVKELEVAIRNRDRLRFGLYSLFEW
FHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY
NESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI
SDYLTIEELVKQLVETVSCGNNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT
WRSQNDTVPDVMYTSKPKKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI
SLEQNGIMVELPQLTIHQMPCKKGWALALTNVI

Signal sequence:

amino acids 1-28

N-glycosylation site.

amino acids 171-175, 239-243, 377-381

Casein kinase II phosphorylation site.

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319,
375-375

Tyrosine kinase phosphorylation site.

amino acids 361-369, 389-397

N-myristoylation site.

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

Leucine zipper pattern.

amino acids 410-432

Alpha-L-fucosidase putative active site.

amino acids 283-295

FIGURE 73

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGT
TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCCTCTCATATCACCAGTGGCCATC
TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTTG
CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCAAGGCTCTTTGCGTGCAAGAAGCT
TTCCATCCAGGTGTCTATGCAGAATTATGGGGATCACCTTTGTGAGCAAAAAGGCGAACCAGC
AGCTGAAATTTACAGTAAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG
GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGTATGGCTGGGTTGGAGA
TGGGATCGTGGTCACTCTAGGATTAGCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTG
TCCTGATTTGGAAAGGTTCCAGTGGAGCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGAT
ACTTGGACTAACTCGTGCAATTCAGAAATTATCACCACCAAGATCCCATATTCAACACTCA
AAGTCAACACAAACACAGAATTTATTGTGAGTACAGTACCTACTCGGTGGCATCCCCTT
ACTCTACAATACCTGCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGG
AGAAAAAATTGATTTGTGTACAGAAGTTTTTATGAAACTAGCACCATTGTCTACAGAAAC
TGAACCATTTGTTGAAATAAAGCAGCATTCAGAAATGAAGCTGCTGGGTTTGGAGGTGTCC
CCACGGCTCTGCTAGTGCTTGTCTCCTCTTCTTGGTGCTGCAGCTGGTCTTGGATTTTGC
TATGTCAAAAGGTATGTGAAGGCCCTTCCCTTTTACAACAAGAAATCAGCAGAAGGAATGAT
CGAAACCAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCTTAATGAGGAATCAAAGA
AAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCGATGCTCGGAA
GCTGAAGTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTTTCATGCTCC
TTACCTTGCCCCAGCTGGGGAATCAAAGGGCCAAAGAACCAAGAAGAAAGTCCACCTT
GGTTCCTTAAGTGAATCAGTCTCAGGACTGCCATTGGACTATGGAGTGACCAAAGAGAATGC
CCTTCTCCTTATTGTAAACCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCACGGCC
TTTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTTGTCAAAGTGCAA
GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAGGCCCTCTGGCTGTCTGAGCTAGG
TGGGTTGAAAGCCAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGAC
CCTTCTTTCAGCTCTGAAAGAGAAACAGTATCCCACTGCATGTCTTCTGAGCCCGGTA
AGAGCAAAAGAAATGGCAGAAAAGTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAG
ACCTAATCTCTGTAAAGCTAAAATAAGAAATAGAACAAGGCTGAGGATACGCAGTACACT
GTCAGCAGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGA
ATCACTGTTTAGAACACACACACTTACTTTTTCTGGTCTCTACCCTGCTGATATTTTCTCT
AGGAAATATACTTTTACAAGTAAACAAAATAAAAACCTCTTATAAATTTCTATTTTATCTGA
GTTACAGAAATGATTACTAAGGAAGATTACTAGTAATTTGTTTTAAAAAGTAATAAAATTCA
ACAAACATTGTGTAATAGCTACTATATGTCAAGTGTGTGCAAGGTATTACACTCTGTAAT
TGAATATTATCTCAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTTCAT
GTTTTGATATTTCTAGCTTATCTACTTCCAACTAATTTTTATTTTTGTGAGACTAATCTT
ATTCAATTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAAACATACCTAAGAG
TACATTGTTACCTCTATATACCAAGCAGATTTTAAAGTGCCATTAAACAAATGTATCACTA
GCCCTCCTTTTTCCCAACAAGAGGACTGAGAGATGCAGAAATATTTGTGACAAAAATTA
AGCATTTAGAAACTT

FIGURE 74

MARCFSLVLLLSIWTTTRLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACR
LLGLSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQF
AAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTPP
APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAPKNEAAGFGGVPTALLVLALLFF
GAAAGLGFCYVKRYVKAPFTNKNQKEMIETKVVKEEKANDSNPNEEKSKTKDNPEESKSP
SKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 75

AGATGCGGCTCTTGGCACCTCTAATTGCTCTCGTGATTTCGGTGCCGCGACTTTCACGATGG
CTCGCCCAACCTTACTACCTTCTGTGCGCCCTGCTCTCTGCTGCCTTCTCTACTCGTGAGGAA
ACTGCCGCGCTCTGCCACGGTCTGCCACCCAAACGGAAGACGGTAACCCGTGTGACTTTG
ACTGGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGA
TCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAAT
TCTTTTCTTCCGCTTGGATATTGCGATGGGCCTACTTTACATCACACTCTGCATAGTGTTC
TGATGACGTGCAAAACCCCCCTATATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAA
ACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACCTGGATTGTGGAGTCTTTTGCCAA
TTGGTCTAATGACTGCCAATCATTTGCCCTTATCTATGCTGACCTCTCCCTTAAATACAACT
GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTAC
AAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCTGATCCTGTTCCAAGGTGGCAA
GGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTG
AGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTATACGACGGGCCAAGAACTATCA
AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCAACCACAGTGTGAGA
TGGGGAAAACAAGAAGGATAAA**TAAG**ATCCTCACTTTGGCAGTGCTTCTCTCTGTCAATT
CCAGGCTCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTNNATNNATGTTTTCCCTTTGG
CTGNGACTGGNTGGGGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTGAG
GCACCCTACAGGAAGGCCTGCCATGCTGTGGCCAACTGTTTCACTGGAGCAAGAAAGAGATC
TCATAGGACGAGGGGGAAATGGTTTCCCTCCAAGCTTGGGTGAGTGTGTTAACTGCTTATC
AGCTATTGAGACATCTCCATGGTTTCTCCATGAACTCTGTGGTTTCACTATTCCTTCTTAG
TTGACCTGCACAGCTTGTTTAGACCTAGATTTAACCCCTAAGGTAAGATGCTGGGGTATAGAA
CGCTAAGAATTTTCCCCAAGGACTCTTGCTTCTCTTAAGCCCTTCTGGCTTCGTTTATGGTC
TTCATTAAAGTATAAGCCTAACTTTGTGCTAGTCCTAAGGAGAAACCTTTAACCCAAAG
TTTTTATCATTGAAGACAATATTGAACAACCCCTATTTTGTGGGGATTGAGAAGGGGTGAA
TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCTGGACTTTCAC
TAACCCCTGACATACTCCCAACCCAGTTGATGGCTTTCCTGTAATAAAAGATTGGGATT
TCCTTTTG

FIGURE 76

MAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAFLLVRLPPLCHGLPTQREDGNPCDFD
WREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVFL
MTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNC
TGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSE
ENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

Signal sequence:

amino acids 1-48

Transmembrane domain:

amino acids 111-125

N-glycosylation site.

amino acids 165-169, 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 154-158, 265-269

Casein kinase II phosphorylation site.

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

N-myristoylation site.

amino acids 188-194, 225-231

Myb DNA-binding domain repeat signature 1.

amino acids 244-253

[illegible][illegible]

FIGURE 78

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNHGKDLLNGVKLVVETPEETLFTYQ
GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLAIGLRHRSFGDYQGRVHLRQD
KEHDVSLEIQDLRLLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQ
VCAEQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRYSYGPR
HRR LHRYDVFCFATALKGRVYYLEHPEKLTLTAREACQEDDATIAKVGQLFAAWKFHGLDR
CDAGWLADGSVRYPVVHPHPNCGPPEPGVRSFGFPDPQSRLYGVCYRQH

Signal sequence:

amino acids 1-17

Casein kinase II phosphorylation site.

amino acids 29-33, 53-57, 111-115, 278-282

Tyrosine kinase phosphorylation site.

amino acids 137-145

N-myristoylation site.

amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

FIGURE 79

GGAGAGCGGAGCGAAGCTGGATAACAGGGGACCGATGATGTGGCGACCATCAGTTCTGCTGC
TTCTGTTGCTACTGAGGCACGGGGCCAGGGGAAGCCATCCCCAGACGCAGGCCCTCATGGC
CAGGGGAGGGTGCACCAGGCGGCCCCCTGAGCGACGCTCCCCATGATGACGCCACAGGGAA
CTTCAGTACGACCATGAGGCTTTCTGGGACGGGAAGTGCCAAAGGAATTCGACCAACTCA
CCCCAGAGGAAAGCCAGGCCCTCTGGGGCGGATCGTGGACCGCATGGACCGCGCGGGGGAC
GGCGACGGCTGGGTGTCTGCTGGCCGAGCTTCGCGCGTGGATCGCGCACACGCAGCAGCGGCA
CATACGGGACTCGGTGAGCGCGGCCCTGGGACACGTACGACACGGACCCGCGACGGGCGTGTGG
GTTGGGAGGAGCTGCGCAACGCCACCTATGGCCACTACGCGCCCGTGAAGAATTTTCATGAC
GTGGAGGATGCAGAGACCTACAAAAAGATGCTGGCTCGGGACGAGCGCGTTTCGCGGTGGC
CGACCAGGATGGGGACTCGATGGCCACTCGAGAGGAGCTGACAGCCTTCCTGCACCCCGAGG
AGTTCCCTCACATGCGGGACATCGTGATTGCTGAAACCTGGAGGACCTGGACAGAAACAAA
GATGGCTATGTCAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCAGCCTGGGGAGGA
GGAGCCCGCGTGGGTGCAGACGGAGAGGCAGCAGTTCCGGGACTTCGGGATCTGAACAAGG
ATGGGCACCTGGATGGGAGTGAGGTGGGCCACTGGGTGCTGCCCTGCCAGGACCAGCCC
CTGGTGGAAGCCAACCCTGCTGCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAA
AGCGGAAATCCTGGGTAATTGGAACATGTTTGTGGGCACTCAGGCCACCAACTATGGCGAGG
ACCTGACCCGGCACCACGATGAGCTGTGAGCACCCGCGCACCTGCCACAGCCTCAGAGGCCCG
CACAATGACCGGAGGAGGGGCCGCTGTGGTCTGGCCCCCTCCCTGTCCAGGCCCCGAGGAG
GCAGATGCAGTCCAGGCATCTCTGCCCTGGGCTCTCAGGGACCCCTGGGTGGCTTC
TGTCCCTGTACACCCCCAACCCAGGGAGGGCTGTCTAGTCCCAGAGGATAAGCAATAC
CTATTCTGACTGAGTCTCCAGCCCAGACCCAGGGACCCCTGGCCCCAAGCTCAGCTCTAA
GAACCGCCCCAACCCCTCCAGTCCAAATCTGAGCCTCCACCACATAGACTGAAACTCCCTT
GGCCCCAGCCCTCTCTGCTGGCCTGGCCTGGGACACCTCCTCTCTGCCAGGAGGCATAAA
AAGCCAGCGCCGGGACCTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 80

MMWRPSVLLLLLLLRHGAQGKPSPDAGPHGQGRVHQAPLSDAPHDDAHGNFYDHEAFLGR
EVAKEFDQLTPESQARLGRIVDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDT
YDTRDRGVRGWEEELRNATYGHYAPGEEFHDVEDAETYYKMLARDERRFRVADQGDGSMATRE
ELTAFLHPPEEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEEPAWVQTERQQ
FRDFRDLNKGHLDGSEVGHVLPAPAQDQPLVEANHLLHESDTRDKDGRLSKAEILGNWNMFV
GSQATNYGEDLTRHHDEL

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 140-144

Casein kinase II phosphorylation site.

amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293,
291-295, 298-302

N-myristoylation site.

amino acids 263-269, 311-317

Endoplasmic reticulum targeting sequence.

amino acids 325-330

FIGURE 81

GGGGCCCTTGCTTCCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGG
 GCGGCGGGCGCGGGTGGAGGGGATCCCTGACGCCTCTGTCCCTGTTCTTTGTGCGTCCCAG
 CCTGTCTGTCTGCTGTTTTGGCGCCCCCGCTCCCGCGGTGCGGGGTGCACACCGATCCTG
 GGCTTCGCTCGAATTGCGCGCAGGCGCCTCCAGACCTAGAGGGGCGTGGCTGGAGCAG
 CGGGTGTCTGTGTCTCTCTCTCTGCGCGCGCCCGGGGATCCGAAGGGTGGGGGCTCT
 GAGGAGGTGACGCGCGGGGCTCCCGCACCTTGGCCTTGCCCGCATTTCTCCCTCTCTCCAG
 GTGTGAGCAGCCTATCAGTCACCAATGTCCGCAGCCTGGATCCCGGCTCTCGGCTCGGTGTG
 TGTCTGTCTGCTGCTGCCGGGGCCCCGGGGCAGCGAGGGAGCCGCTCCCATTGCTATCACATG
 TTTTACCAGAGGCTTGGACATCAGGAAAGAGAAAGCAGATGTCCTCTGCCCAGGGGGCTGCC
 CTCTTGAGGAATTTCTGTGTATGGGAACATAGTATATGCTTCTGTATCAGGCATATGTGGG
 GCTGCTGTCCACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTTAGCCCTACC
 TGGTCGAGAAAACATTCTCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGAT
 GGTCTGCTTCTTTACAGTAACATAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCA
 GTGTCCACAGCACATCCACCAACAGGTAAACGACTAAAGAAAAACCCCGAGAAGAAAACTGG
 CAATAAAGATTGTAAAGCAGACATTGCATTCTTGATTGATGGAAGCTTTAATATTGGGCAGC
 GCCGATTTAATTACAGAAGAAATTTGTTGGAAAAGTGGCTCTAATGTTGGGAATTGGAACA
 GAAGGACCCATGTGGGCTTGTTCAGGCCAGTGAAACATCCCAAAATAGAATTTTACTTGAA
 AAACCTTTACATCAGCCAAAGATGTTTTGTTTGCCATAAAGGAAGTAGGTTTCAGAGGGGGTA
 ATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAAATTTCTCAGGTAGATGCTGGA
 GTAAGAAAAGGGATCCCAAGTGGTGGTGGTATTTATTGATGGTTGGCCTTCTGATGACAT
 CGAGGAAGCAGGCATTGTGGCCAGAGAGTTTGGTGTCAATGTATTATAGTTTTCTGTGGCCA
 AGCCTATCCCTGAAGAACTGGGGATGGTTCAGGATGTACATTTGTTGACAAGGCTGTCTGT
 CGGAATAATGGCTTCTTCTCTTACCACATGCCCAACTGGTTTGGCACCACAAAATACGTAA
 GCCTCTGGTACAGAAGCTGTGCACCTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACT
 CAGTGAACATTGGCTTTCTAATTGATGGCTCCAGCAGTGTGGAGATAGCAATTTCCGCTTC
 ATGCTTGAATTTGTTTCCAAATAGCCAAGACTTTTGAAATCTCGGACATTGGTGCCAAGAT
 AGCTCTGTACAGTTTTACTTTATGATCAGCGCACGGAGTTTCAGTTTCACTGACTATAGCACC
 AAGAGAATGTCTAGCTGTCTATCAGAAACATCCGCTATATGAGTGGTGGAAACAGCTACTGGT
 GATGCCATTTCCCTTCACTGTTAGAAATGTTTTGGCCCTATAAGGGAGAGCCCCAACAGAA
 CTTCTAGTAATTGTACAGATGGGCAGTCCTATGATGATGTCCAAGGCCCTGCAGCTGCTG
 CACATGATCAGGAATCACTATCTTCTCTGTTGGTGTGGCTTGGGCACCTCTGGATGACCTG
 AAAGATATGGCTTCTAAACCGAAGGAGTCTCAGCCTTTCTCACAAGAGAGTTTCAAGGATT
 AGAACCAATTGTTTCTGATGTCTATCAGAGGCATTTGTAGAGATTTCTTAGAATCCCAGCAAT
AATGGTAACATTTTGACAACCTGAAAGAAAAGTACAAGGGGATCCAGTGTGTAATTTGAT
 CTCATAAATACTGAAATGCTTTAGCATACTAGAATCAGATACAAACTATTAAGTATGTCAAC
 AGCCATTTAGGCAAATAGCACTCCTTTAAAGCCGCTGCCCTCTGGTTACAATTTACAGTGT
 ACTTTGTTAAAAACACTGTGAGGCTTCAATCATGGCTCTTAGAACTCAGGAAAGAGGA
 GATAATGTGGATTAAACCTTAAGAGTTCTAACCACTCTAATATGTACAGATATGCAAA
 TTCCATAGCTCAATAAAGAATCTGATACTTAGACCAAAAAAAAAA

FIGURE 82

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCTFTRGLDIRKEKADVLCPGGCPLEEFVS
YGNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASFTVT
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDCKADIAFLIDGSFNIGQRRFNLQKN
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLNFTSAKDVLFAlKEVGFGRGNSNTGKAL
KHTAQKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELG
MVQDVTFVDKAVCRNNGFFSYHMPNWFGTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIAFLI
DGSSSVGDSNFRMLLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVI
RNIRYMSGGTATGDAISFTVRNVFGPIRES PNKNFLVIVTDGQSYDDVQGPAAAHADAGITI
FSVGVAWAPLDDLDKMASKPKESHAFFTREFTGLEPIVSDVIRGICRDFLESQQ

Signal sequence:

amino acids 1-24

N-glycosylation site.

amino acids 100-104, 221-225

Casein kinase II phosphorylation site.

amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424,
425-429, 478-482, 528-532

N-myristoylation site.

amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211,
239-245, 240-246, 261-267, 403-409, 442-448, 443-449

Amidation site.

amino acids 145-149

FIGURE 83

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGCGAGCCC
GGCGGCTCCCGCGGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCAGTCG
GGCGCGCGGCTGCGGGCGCAGAGCGGAGATGCAGCGGCTTGGGGCCACCTGCTGTGCTGCTG
TGCTGGCGCGCGGCTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTC
AAGCCCGGCGCGGCTCTCAGCTACCCGAGGAGGAGGCCACCTCAATGAGATGTTCCGCGA
GGTTGAGGAACGTATGGAGGACACGACGACAAATTGCGCAGCGCGGTGGAAGAGATGGAGG
CAGAAGAAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCAGCTAT
CACAATGAGACCAACACAGACACGAAGGTTGGAATAATACCATCCATGTGCACCGAGAAAT
TCACAAGATAACCAACAACCAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCATATCTG
TGGGAGACGAAGAAGGCAGAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCAGC
ATGTACTGCCAGTTTGCCAGCTTCCAGTACACCTGCGAGCATGCGGGGCCAGAGGATGCT
CTGCACCCGGGACAGTGAGTGCTGTGGAGACCAGCTGTGTGCTGTGGGTCTACATGCACCAAAA
TGGCCACAGGGGCGCAATGGGACCATCTGTGACAAACGAGGGGACTGCCAGCCGGGGCTG
TGCTGTGCCCTTCCAGAGAGGCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGA
GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATACCTGGGAGCTAGAGCCTGATG
GAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTG
GTGTATGTGTGCAAGCCGACCTTCGTGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCC
CAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG
AGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTCGCGCTGCCCCGCTG
GCACTGCTGGGAGGGGAAGAGATT**TAG**ACTGTGGACCAGGCTGTGGGTAGATGTGCAATAGAA
ATAGCTAATTTATTTCCCGAGGTGTGTGCTTTAGGCCGTGGGCTGACCAGGCTTCTTCTCTACA
TCTTCTTCCCGAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTTCAGC
TCCCCCAGGCTGTTCTCCAGGCTTACAGTCTGGTGCTTGGGAGAGTCAAGCAGGGTTAAAC
TGCAGGAGCAGTTTGCCACCCCTGTCCAGATTATTGGTGCTTTGCTCTACAGTTGGCAG
ACAGCCGTTTGTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAACAAATGTGG
AGTCTCCTCTGATTGGTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAAACATCAA
CCTGGCAAAAATGCAACAAATGAATTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTG
TGCCTTCAGCTGTTGCAGATGAATGTTCTGTTTACCCTGCATTACATGTGTTTATTTCATCC
AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTC
CCTCTCTCAGCACAGCCTGGGGAGGGGGTCAATTGTTCTCCTCGTCCATCAGGGATCTCAGAG
GCTCAGAGACTGCAAGCTGCTTGCCCAAGTCAACAGCTAGTGAGAACCCAGAGCAGTTTCAT
CTGGTTGTGACTCTAAGCTCAGTGCTCTCTCCACTCCCAACAGCAGCCTTGGTGCCACCAA
AAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTCTTGAGGCATGCACATCTGGAATTAAG
GTCAAACTAATTTCTCACATCCCTCTAAAAGTAACTACTGTTAGGAACAGCAGTGTTCTCAC
AGTGTGGGGGAGCCGCTCTTCTAATGAAGACAATGATATTGACACTGCTTCTTTGGCAGT
TGCTATTAGTAACTTTGAAGGTATATGACTGAGCGTAGCATACAGGTTAACCTGCAGAAACA
GTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGCAAAATCACTTAGCAGCAAC
TGAAGACAATTTACAAACACGTGGAGAAAATCAAACGAGCAGGGCTGTGTGAACATGGTT
GTAATATTGCGATCTGCAACACTGAACTCTACGCCATCCCAAAATGATGTTTTCAGGTGTCA
TGGACTGTGCCACCATGTATTATCCAGAGTTCTTAAAGTTTAAAGTTGCACATGATTGTA
TAAGCATGCTTTCTTTGAGTTTAAATTTATGTATAACATAAGTTGCATTTAGAAATCAAGC
ATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 84

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYQEEATLNEMFREVEELMEDTQ
HKLRSAVEEMEAEEAAKASSEVNLANLPSPYHNETNTDTKVGNNTIHVHREIHKITNNQTG
QMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQRMCLTRDSECCG
DQLCVWGHCTKMATRGSGNTICDNQRDCQPGLCACAFQRLFPVCTPLPVEGELCHDPASRL
LDLITWELEPDGALDRCPASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV
GSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEET

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 85

AAGGAGGCTGGGAGGAAAGAGGTAAGAAAGGTTAGAGAACCTACCTCACAATCTCTCTGGGCTCAGAAGGACTCTG
AAGATAACATAATTTCCAGCCCTCCACTCTCCTCCCTCCCAACACACATGTGCATGTACACACACATACAC
CACACATACACCTTCTCTCTCTCACTGAAGACTCACAGTCACCTCACTCTGTGAGCAGCTCATAGAAGAGACAC
TAAAGCCTTAAGAGCAGSCCTGGCCATTAACCTCTGCAGCTCCGTTTGGCTGTGTGATCAAAAACATGGGAGGGG
CCAGGCACGGTGACTCACACCTGTGAATCCAGCATTCTGGAGACCGAGGTCAGGCAGATCACTTGAGGTTCAGGAG
TTCGAGACCCAGCCTTGCCCAACATGGAGAAACCCCATCTCTACTAAAAATACAAAAATTAGCCAGGAGTGGTGGC
AGGTGCTGTAAATCCAGCTACTCAGGTGGCTGAGCCAGGAGAATCGCTTGAATCCAGGAGGCCAGGAGTGCAGT
CAGCTGAGTGACCGCTGCACCTCCAGCCTGGGTGACAGAATGAGACTCTGTCTCAAAACAAACAAACCCGGAGGA
GGGGTAGATATGTCTCTCTGCAACCTCTTAACTCTGCATCTCTCTTCCAGGCTGCCCTTGATGGGGCTCCTG
GCAATGACTGAGCAGGCCAGCCCCAGAGGACAAGGAAGAGAAGGCAATTATGGAGGAGGCAAGAAGTGACGCCCG
GTGTGAATGACTGCCCTGGGAGGTTGGTTCTTGGGCCCTGGCAGGGTTGCTGAACCTTACCTGCAAAAACACA
AAGAGCAGGACTCCAGACTCTCCTTGTGAATGGTCCCTGCCCTGCAGCTCCACCAATGAGGCTTCTCTGTGGCCCC
ACTCTTGCTAGCTTGGGTGGCTGGTGCCACTGCGCACTGTGCCGTGGTACCTGGCATGTTCCCTGCCCCCTCA
GTGTGCTGCCAGATCCGGCCCTGGTATACGCCCCCGCTCGTCTACCGCAGGGCTACCACTGTGGACTGCAATGA
CCTATTCTGACGGCAGTCCCCCGGCACTCCCCGAGGCACACAGACCCTGCTCTGTGAGAGCAACAGCATTGT
CCGTGTGAGCAGAGGTGAGCTGGGCTACCTGGCCAAATCTCACAGAGCTGGAGCTGTCCCAACAGACTTTTCGGA
CTCTCCGAGACTGTGATTTCATGCCCCAGCTGTGAGCTGTGAGCTGACCTAGAGGAGCAACAGCTGACCCGGCT
GGAGGACCAAGCTTTGCGAGGCTGGCCAGCCTACAGGAACCTCTATCTCAACCAACACAGCTCTACCGCATCGC
CCCCAGGCGCTTTTCTGGCTCTCAGCAACTGTCTGAGCTGCACCTCAACTCACTCTCTGAGGCGACTTGACAG
CCGCTGGTTGAAATGTGCCCACTTGGAGATCTCATGATTGGCGGCAACAGGTAGATGGCATCTCTGGACAT
GAATCTCTCGGCCCTGCGCACTCTGCTAGCTGGTGTCTAGCAGGCATGAACCTGGCGGAGATCTCCGACTATGC
CCTGGAGGGGCTGCAAGAGCTGGAGAGCTCTCTCTTCTATGACAACAGCTGGCCCGGGTGCCGAGGCGGGCACT
GGAAAGGTGCCCGGGCTCAAGTTCCTAGACCTCAACAAGAACCCGCTCCAGCGGGTAGGGCGGGGGACTTTGC
CAACATGCTGCACCTTAAGGAGCTGGGACTGAACAACATGGAGAGCTGGTCTCCATGCACAAGTTTGGCTTGGT
GAGCTTCCCCGAGCTGACCAAGCTGGACATCAACCAATAACCCAGCGCTGCTCTTCACTCACCCCGGCCCTTCCA
CCACTTGCACCAGATGGAGACCTCTATGCTCAACAACACAGCTCTCAGTGCTCTGATCAGCAGGACAAGTGAGTG
CTCGCCCAACCTGCAGGAGTGGTCTTCCAGGCAACCCCATCCGCTGTGCTCACTCCCAACCCCGGCCCTTCCA
CACGGCAGCCGTGTCTCGTCTCATGAGCGCAATCCACCTGTGTGCGAGCTCTCCGAGCTCCAGCGCTCTCC
GGTCCGTGAGTGGCTTCCGAGTACGACGAGCACTGTTTGGCCCTCATCTCCCAACCCCGGCCCTTCCCCAAG
CCTCCAGGTAGCCAGTGGAGAGAGCATGGTGTGCTGATGTCGCGGCACTGGCCGAAACCCGAGAGTCTACTG
GGTCACTCCAGCTGGGCTTGCATGACACCTGCCCATGCAAGGAGGATACCGGTGTACCCGGGGGAGCCTT
GGAGCTGCGAGGGGTGACAGCAGAAGGCGAGGCTATACACCTGTGTGGCCAGAACCTCTGGGAGGCTGACAC
TAAGACGGTTAGTGTGTTGTGGGCCGTCTCTCTCCAGCCAGGCGAGGACGAAGGACAGGGGCTGGAGCTCCG
GGTGAGGAGACCACTCCCTATTCACATCTCTGCTATCTTGGGTCAACCCCAACACAGATGTGCCAACCACTCAC
CTGGTCCAGTGCTCTCTCCTCCGGGCCAGGGGGCCACAGCTCTGGCCCCGCTGCTCTCGGGAAACCCACAGCTA
CAACATTACCGCCTCTCTTCAGGGCCACGAGTACTGGGCTGCTGTGCAAGTGGCCTTTGCTGATGCCCAACCA
GTTGGCTGTGTATGGGCGAGGACCAAGAGGCCACTTCTTGCCACAGAGCCTTAGGGGATCGTCTCTGGGCTCAT
TGGCATCTGCTCTCTGCTCTTCTCTGCGCAGCTGGGCTAGCGGCCCACTTGGCACAGGCCCAACCCAGGAA
GGGTGTGGGTGGGAGGCGGCTCTCCTCCAGCCTGGGCTTCTGCGGCTGGAGTGCCCTTCTGTCTCGGGTTGT
GTCTGCTCCCTCTCTCCTGCCCTGCGAATCCAGGAGGAGTCTGCCAGATCTCTAGAAGGGGAGACACTGTTGCC
ACCATGTCTCAAAATTTCTGAAGACTCAGCTGTTTCTCAGCAGTAGAGAAATCACTAGGACTACTTTTACCAA
AGAAAGCAGCTTGGGCCAGATGCCCCGCCAGGAAGGACATGGACCCACGCTTGAAGCCTTGCAGCTGGGC
CAAGAAGATGGGGCTTTGTGGCCCTGGGGGTGCTTCTGCAGCCTTGAAGAAAGTGCCCTTACCTCTCAGGCTCA
CCTCTGCTGCATTCTGAGGAACATCTCCAAGGAACAGGAGGACTTTGGCTGAGGCGCTCTGCTCTCCCACTTCT
CTCTCTGCGAGAGGCTCTGGGCTGGCTTGGCTGTCCCTACTGTGTCCCAGGGCTGACCCCTTCTCTCTC
TCTTTCTGTACAGTCTCAGTTGCTTGTCTTGTGCTCTGGGCAAGGGCTGAAGGAGGCTCATCATCTCAC
CTCGGGGGCTGCCTCAATGTGGGAGTGACCCAGCCAGATCTGAAGGACATTTGGGAGAGGATGCCAGGAA
CGCCTCATCTCAGCAGCTGGGCTGGCATTCGAAGCTGACTTTCTATAGGCAATTTGTGACATTTGTGGAGAA
ATGTGTCACTTCCCCAACCCGATTCACTCTTTCTCTCTTTTGTAAAAAATAAAAAATAATAACAAATAAA
AAAA

FIGURE 86

MRLVAPLLLLAWAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVCNDLFLTA
VPPALPAGTQTLLQSNISIVRVDQSELGYLANLTELDLSQNSFSDARDCDFHALPQLLSLHL
EENQLTRLEDHSFAGLASLQELYLNHNQLYRIAPRAFSGLSNLLRLHLNSNLLRAIDSRWFE
MLPNEILMIGGNKVDAILDNMNFRPLANLRSVLVLAGMNLREISDYALEGLQSLESLSFYDNO
LARVPRRALEQVPGLKFLDLNKNPLQRVGPGDFANMLHLKELGLNNMEELVSIDKFALVNLPL
ELTKLDITNNPRLSFIHPRAFHHLPMETMLMNNALSALHQQTVESLPLNQEVGLHGNPIR
CDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLPVREVPFREMTDHLPLISPRSFPPSLQ
VASGESMVLHCRALAEPEPEIYWVTPAGLRITPAHAGRRYRVYPEGTLELRRVTAEAGLYT
CVAQNVLVADTKTVSVVVGRALLQPRDEGQGLELRVQETHPYHILLSWVTPPNTVSTNLTW
SSASSLRGQATALARLPRGTHSYNITRLLQATEYWACLQVAFADAHTQLACVWARTKEATS
CHRALGDRPGLIAILALAVLLAAGLAHLGTGQPRKGVGGRRPLPPAWAFWGSAPSVRVW
SAPLVLPWNPRKLPSSSEGETLLPPLSQNS

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 629-648

N-glycosylation site.

amino acids 94-98, 381-385, 555-559, 583-587

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 485-489

Casein kinase II phosphorylation site.

amino acids 46-50, 51-55, 96-100, 104-108, 130-134, 142-146,
243-247, 313-317, 488-492, 700-704

Tyrosine kinase phosphorylation site.

amino acids 532-540

N-myristoylation site.

amino acids 15-21, 493-499, 566-572

Amidation site.

amino acids 470-474, 660-664, 692-696

FIGURE 87

GCAAGCCAAAGCGCTGTTTGAAGAAGTGAAGAAGTTCGGACCCATGTGGAGGAGGGGGACATTGTGTACCGCCT
 CTACATGCGGAGACCATATCAAGGTGATCAAGTTTCATCTCATCTGCTACACCGTCTACTAGCTGACAA
 CATCAAGTTGACAGTGGAGCTGACCGTGGACATTTAGAGCCTGACGGCTACCGACCTTACCGCTGTGCCACCC
 CCTGGCCACATCTTCAAGATCCTGGCGTCTTTACATCAGCCTTAGTCACTCTTACGCGCTCATCTGCATGTA
 CACACTGTGTGGATGCTACGCGCTCCCTCAAGAAGTACTGTTTGAGTCGATCCGTTGAGGAGGACGATCAG
 CGACATCCCCGACGCTCAAGAACGACTTCGCTTCATGCTGCACCTCATTGACCAATACGACCGCGCTTACTCCAA
 GCGCTTCGCGCTTCTTCTGCGGAGGTGAGTGAGAACAAAGCTGCGGACGTTGAACCTCAACCAACGAGTGGACGCT
 GGAACAAGCTTCGCGACGCGCTCACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTTGTTCATGCTCAGTGGCAT
 CCTGTACATCTGTTTGACTCTGTTGGAGCTGGAGGTCTCAAGCTGGAGCTGATCCCGAGCTGACCATCCGCGC
 CAGCATTTGCCAGCTCAGCGGCTCAAGGAGCTGTGGCTCTACACACAGCGCGCAAGATTGAAGGCGCTTGGCT
 GGCCTCTCTGCGGAGAGAACTGCGGGCGCTGCACATCAAGTTCAACGACATCAAGAGATCCCGCTGTGGATCTA
 TAGCTGAAGACACTGGAGGAGCTGCACCTGACGGGCAACTGAGCGGGAGAACAAACCGCTACATCGTCATCGA
 CGGGCTGCGGAGCTCAACAGCCTCAAGGTGCTGCGGCTCAAGAGCAACCTTAAGCAAGCTGCCACAGTGGTCAC
 AGATGTGGGCTGCACCTGCAGAACTGTCCATCAACAATGAGGGCACCAGCTCATCGTCTCAACAGCCTCAA
 GAAGATGGCGAACTGACTGAGCTGGAGCTGATCCGCTGCGACCTGGAGCGCATCCCCACCTTCATCTTCAGCCT
 CCACAACCTGCAGGAGATTGACCTCAAGGACAACAACTCAAGACCATCAGGAGATCATCAGCTTCCAGCACCT
 GCACCGCTCACCTGCTTAAGCTGTGTTACAAACCATCGCTTACATCCCCCTCAGATCGGCAACCTCAACAA
 CCTGGAGCGCTCTAATGAAACCGCAACAGATCGAGAAGATCCCCACCCAGCTCTTCTAGTGGCGGCAAGCTGCG
 CTACTCTGGACCTCAGCGCAACAACTGACCTTCTCTCTCGGACATCGGCTCTCTCGAGAACTCCAGAACTCT
 AGCCATCACGCTGACCGATCGAGCGCTCCTCCGAGCTCTTCCAGTCCCGAAGCTGCGGGCTGACCTGACCT
 GGGCAACACGCTGCTCCGCTCAGCTGACCTTCTCTCTCGGACATCGGCTCTGAGCAGATCGAGCTGCGGGG
 CAACCGGCTGAGTGCCTGCTGTGGAGCTGGCGGAGTGCCTCACTCAAGCGCAGCGCTTGGTGTGGAGGA
 GGACTCTTTCAACACATCGCCCAACCGAGGTGAAGGAGCGGCTGTGAGGCTGACAGCAGATCGAGCTGCGGG
 GCGCGCCAGCAAGCAAGCAGCAGGACCGCTGCCAGTCTCTCAGCGCGGAGGCGGAGCGCTAGCTTCTCCAG
 AACTCCGCGCAGCAGCAGCAGCCTCGCGCTGGCGAGGAGCTGGGGCTGCTTGTAGTCAAGCGCAGAGCGAGA
 GGAAGTATCTGTGGGCTGGCGCTTTCTCCCTCTGAGACTCAGCTCCCGAGGGCAAGTGTCTGTGGAGGAG
 AGCAAGTCTCAAGAGCGCGATTTGGATAATCAGGGTCTCTCTCTGGAGGCGAGCTCTGCCACAGGCTGAG
 CTGCGCACAGAGCTGCGGACCTCACTTTAGTTCTTGGTATTTATTTTCTCCATCTCCCACTCTTTCATCC
 AGATAACTTATACATTTCCAAAGAAAGTTTCAAGCCAGATGGAAAGTGTTCAGGAAAGGTTGGGCTGCTCTTCTCC
 TTTGCTTATTATGAGCATGCCGCGGGCATTTAAACCCACCTGGAATTCAGCAGAGTGTCTCGGGGCGAACAG
 CCAATGGGACGGTCAACCAAGCTGCGGGCTGGGCTTGCAGTTCAGGAGTGTGCGGAGAGCAGGCTCCAGCTGGA
 AAGGCCAGGCGTGGAGCTTGCCTCTTCAGTTTGTGGCAGTTTGTAGTTTGTGTTTTTTTAAATCAAA
 AAACAATTTTTTTTAAAAAAGCTTTGAAAAATGGATGTTTGGGTATTTAAAAAGAAAAAATCTTAAAAA
 AAAAGACACTAACCGCCAGTGTGAGTTGGAGTCTCAGGCGAGGTTGGCAGTTTCCCTTGAGCAAGCAGCCAGCT
 TGAATGTGTTTCTTCTTCTGGGCGAGGTTGACGGGTGCTTCCGATCTGCTGTGACCTTGGTCCAGGAGTT
 CTATTTGTCTTGGGGAGGAGTTTGTGTTTGTGTTTTTGGGTTTTTTTGGTGTCTGTTGTTTCTCTCTCT
 ATGTGTTCTGGCAGGCACTTCTTCTGTGCTGTGCGCAGAGGGAATGTTCTGAGAGCTGCCAAGGAGGAGGAG
 ACTCGGTTGGCTAATCCCGGATGAACGGTGTCTTCCATCGCACCTTCCCTCTCTGCTGCTGCTGCTCTCA
 CGCACAGTGTAAAGAGCCAAAGAGGACCACTTCGCCAGACTTGTGTTCCCACTCTCTGCGCATGGGTGTGT
 CCACTGCAACCGCTGCGCTCGCTTCCATCAGCCCTGTGCGCACTGGTCTTCAAGAAGAGCAGACCTTA
 GAGGCTGTTGGGAATGGGAGGTTGCGCCTGGGAGGCGAGGCGTGTGTTCCAAGCCGTTCCCGTCCCTGCGC
 CTGGAGTCACACAGCCCACTGCGCACTGTGGTGGCTGGAAGCAACCTGTATGATCACTCGGGTCCCGCCT
 AGAAGGTTCCCGCTTAGATCAATCAGTGACACTAAGGCAAGTTTATAGTCTCTGTCTTAATGATTAATG
 CCATCGCTGTGCTGCTTGTGTTTCTGCTGCTGATTTGATATATCTCAGAAATATGTCACATGAC
 CCTCTGACAAACATGAAGCAAAATTCGTTACATGTGGGTCTGAACCTGTAGCTCGGTACAGTATCAATAAA
 ATCTATAACAGAAAAAATAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 88

MRQTIIKVIKFIILICYTVVYVHNIKFDVDCTVDIESLTGYRTYRCAHPLATLFKILASFYI
SLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPVKNDFAFMLHLIDQYDPLYSK
RFAVFLSEVSENKLRQLNLNNEWTLDKLRQLTKNAQDKLELHLFMLSIGIPDTVFDLVELEV
LKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFRENLRALHIKFTDIKEIPLWI
YSLKTLEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSI
NNEGTKLIVLNSLKKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDNNLKTIIEEISFQ
HLHRLTCLKLWYNHIAIYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKRLYDLSSHNNLT
FLPADIGLLQNLQNLAITANRIETLPPELFQCRKLRLHLGNNVLQSLPSRVGELTNLTQIE
LRGNRLECLFVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

Transmembrane domain:

amino acids 51-75 (type II)

N-glycosylation site.

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 85-89

Casein kinase II phosphorylation site.

amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368,
398-402, 493-497

N-myristoylation site.

amino acids 173-179, 261-267, 395-401, 441-447

FIGURE 89

GCCTGTTGCTGATGCTGCCGTGCGTACTTGTCATGGGAGCTGGCACTGCGGCGCTCTCCCGT
CCCGCGGTGGTTGCTGCTGCTGCCGTGCTGCTGGGCTGAACGCAGGAGCTGTCAATTGACT
GGCCCCACAGAGGAGGGCAAGGAAGTATGGGATTATGTACGGTCCGCAAGGATGCCACATG
TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCTTGTCAT
GTGGCTTCAGGGCGGTCCAGGCGGTCTAGCACTGGATTGGAACCTTTGAGGAAATTTGGGC
CCCTTGACAGTGATCTCAAACCACGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCATTATT
GTGGATAATCCCGTGGGCAGTGGGTTCAGTTATGTGAATGGTAGTGGTGCCATATGCCAAGGA
CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAG
AATTCCAGACAGTTCATTCTACATTTTCTCAGAGTCTATGGAGGAAAAATGGCAGCTGGC
ATTGGTCTAGAGCTTTATAAGGCCATTAGCGAGGGACCATCAAGTGCAACTTTGCGGGGGT
TGCCTTGGGTGATTCTGGATCTCCCTGTTGATTGGTGCTCTCCTGGGGACCTTACCTGT
ACAGCATGTCTCTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTA
CTGAATGCCGTAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT
GATCATTTGAACAGAACACAGATGGGGTGAACCTCTATAACATCTTAACATAAAGCACTCCCA
CGTCTACAATGGAGTCGAGTCTAGAATTACACAGAGCCACCTAGTTTGTCTTTGTGACGCGC
CACGTGAGACACCTACAACGAGATGCCCTAAGCCAGCTCATGAATGGCCCCATCAGAAAGAA
GCTCAAAATTATCTCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGCTCTTTGTGAACA
TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATC
AACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATACCATGGGTGAGGAGGCCTG
GGTGCGGAAACTGAAGTGGCCAGAAGTGCCTAAATTCAAGTCAAGTGAAGTGAAGGCCTGT
ACAGTGACCCTAAATCTTTGGAACATCTGCTTTTGTCAAGTCCACAGAACCTTGCTTTTCT
TACTGGATTCTGAAAGCTGGTCATATGGTTCCTTCTGACCAAGGGGACATGGCTCTGAAGAT
GATGAGACTGGTGACTCAGCAAGAATAGGATGGATGGGGCTGGAGATGAGCTGGTTTGGCCT
TGGGGCACAGAGCTGAGCTGAGGCCGTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT
AACTGGGGCTGTGATCAAGAAGGTCTGACCAGCTTCTGCAGAGGATAAAATCATTGTCTCT
GGAGGCAATTGGAAATTATTTCTGCTCTTAAAAAAACCTAAGATTTTAAAAAATTGAT
TTGTTTTGATCAAAATAAAGGATGATAATAGATATTA

FIGURE 90

MELALRRSPVPRWLLLLPLLLGLNAGVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATNSC
KNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSDLKPRKTTWLQAASLLFVDNPVGTGFSY
VNGSGAYAKDLAMVASDMMVLLKTFFSCHKEFQTVPFYIFSESYGGKMAAGIGLELYKAIQR
GTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVNGLYRE
ATELWGKAEMIIIEQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCLCQRHVRHLQRDAL
QLMNGPIRKKLKIIPEDQSWGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDL
IVDTMGQEAWVRKLKWPPELPKFSQLKWKALYSDPKSLETSAFVKSYKNLAFYWILKAGHMVP
SDQGDMAKMMRLVTQQE

Signal sequence:

amino acids 1-25

N-glycosylation site.

amino acids 64-68, 126-130, 362-366

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 101-105

Casein kinase II phosphorylation site.

amino acids 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

N-myristoylation site.

amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175,
187-193, 195-201, 331-337, 332-338, 360-366

FIGURE 91

GGCCGCGGGAGAGGAGGCCATGGGCGCGCGGGGCGCTGCTGCTGGCGCTGCTGCTGGCTC
GGGCTGGACTCAGGAAGCCGGAGTTCGAGGAGGCGCGCCGTTATCAGGACCATGCGGCCGA
CGGGTCATCACGTCGCGCATCGTGGGTGGAGAGGACGCCGAACTCGGGCGTTGGCCGTGGCA
GGGAGCCTCGCGCTGTGGGATTCCACGTATGCGGAGTGAGCCTGCTCAGCCACCCTGGG
CACTCAGCGCGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATG
GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCTACTACACCCG
TTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTACCCCTATGACATTG
CCTTGGTGAAGCTGTCTGCACCTGTCACCTACACTAAACACATCCAGCCCATCTGTCTCCAG
GCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAA
AGAGGATGAGGCACTGCCATCTCCCCACACCTCCAGGAAGTTCAGGTCGCCATCATAAACA
ACTCTATGTGCAACCACCTCTTCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATG
GTTTGTGCTGGCAACGCCCAAGGCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGACCCTT
GGCCTGTAACAAGAATGGA CTGTGGTATCAGATTGGAGTCGTGAGCTGGGAGTGGGCTGTG
GTGCGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCATTGAGTGGAATCCAGAAG
CTGATGGCCAGAGTGGCATGTCCCAGCCAGACCCCTCCTGGCCACTACTCTTTTCCCTCT
TCTCTGGGCTCTCCCACTCCTGGGCGCGTCTTGAGCCTACCTGAGCCCATGCAGCCTGGGGC
CACTGCCAAGTCAGGCCCTGGTTCTCTTCTGTCTTGTGTTGGTAATAAACACATTCCAGTTGA
TGCCTTGCAGGGCATTTCTTCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 92

MGARGALLLALLLARAGLRKPESQEAPLSGPCGRRVITSRIVGGEDAELGRWFPWQGSRLRW
DSHVCVSVLLSHRWALTAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYFVSNI
YLSPRYLGNSPYDIALVKLSAPVTYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP
SPHTLQEVQVAI INNSMCNHLFLKYSFRKDI FGDMVCAGNAQGGKDACFGDSGGPLACNKNG
LWYQIGVVSWSVGCGRPNRPGVYTNISHHFEWIQKLMAQSGMSQPDPSWPLLFFPLLWALPL
LGPV

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 167-171, 200-204, 273-277

Casein kinase II phosphorylation site.

amino acids 86-90, 134-138, 161-165, 190-194, 291-295

N-myristoylation site.

amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245,
259-265, 269-275

Amidation site.

amino acids 33-37

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 252-263,

Serine proteases, trypsin family, histidine active site.

amino acids 78-84

FIGURE 93

CCCACGCGTCCGCGGACGCGTGGGAAGGGCAGAAATGGGACTCCAAGCCTGCCTCCTAGGGCT
CTTTGCCCTCATCCTCTCTGGCAAATGCAGTTACAGCCCGGAGCCCGACAGCGGAGGACGC
TGCCCCCAGGCTGGGTGTCCCTGGGCCGTGCGGACCCGTGAGGAAGAGCTGAGTCTCACCTTT
GCCCTGAGACAGCAGAATGTGGAAGACTCTCGGAGCTGGTGCAGGCTGTGTGGATCCCAG
CTCTCCTCAATACGGAATAACCTGACCCTAGAGAATGTGGCTGATCTGGTGAGGCCATCCC
CACTGACCTCCACACGGTGCAGAAATGGCTCTTGGCAGCCGGAGCCAGAAAGTGCCATTCT
GTGATCACACAGGACTTCTGACTGTCTGGCTGAGCATCCGACAAGCAGAGCTGCTGCTCCC
TGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGGAACCCATGTTGTAAGGTCCCCAC
ATCCCTACCAGCTTCCACAGGCCCTTGGCCCCCATGTGGACTTTGTGGGGGGACTGCACCGT
TTTCCCCAACATCATCCTCTGAGGCAACGTCTGAGCCGCAAGGTGACAGGGACTGTAGGCCCT
GCATCTGGGGTAACCCCTCTGTGATCCGTAAAGCGATACAACCTTGACCTACAAGACGTGG
GCTCTGGCACCAGCAATAACAGCCAAGCCTGTGCCAGTTCTCGGAGCAGTAATTTCCATGAC
TCAGACCTGGCTCAGTTTCATGCGCTCTTCGGTGCCAACTTTGCACATCAGGCATCAGTAGC
CCGTGTGGTTGGACAACAGGGCCGGGGCCGGGCCGGGATTTAGGCCAGTCTAGATGTGCAGT
ACCTGATGAGTGCTGGTGCCAACTCTCCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAG
GGACAGGAGCCCTTCTGCACTGGCTCATGTGCTCAGTAATGAGTCAGCCCTGCCACATGT
GCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCTACATCCAGCGGGTCA
ACACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCTGCTCTTCGCTCAGGTCAGAGT
GGGGCCGGGTGTTGGTCTGTCTCTGGAAGACACCAAGTTCCGCCCTACCTTCCCTGCCTCAG
CCCCATGTGACCACAGTGGGAGGCACATCCTTCCAGGAACCTTTCCTCATCAGAAATGAAA
TTGTTGACTATATCAGTGGTGGTGGCTTCAGCAATGTGTTCCACGGCCCTTCATACAGGAG
GAAGCTGTAAACGAAGTTCCCTGAGCTCTAGCCCCACCTGCCACCATCCAGTTACTTCAATGC
CAGTGGCCGTGCCTACCCAGATGTGGCTGCACCTTCTGATGGCTACTGGGTGGTCAGCAACA
GAGTGCCCATTCATGGGTGTCCGGAACCTCGGCCCTCTACTCCAGTGTGTTGGGGGGATCCTA
TCCTTGATCAATGAGCAGGATCCTTAGTGGCCGCCCCCTCTTGGCTTTCTCAACCCAG
GCTCTACAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGTGGCTGCCATGAGTCTCTGC
TGGATGAAGAGGTAGAGGGCCAGGGTTCTGTCTCTGGTCTGGCTGGGATCCTGTAAACAGGC
TGGGGAACCAACACTTCCAGCTTTGCTTGAAGACTCTACTCAACCCCTGACCCTTTCCTATC
AGGAGAGATGGCTTGTCCCTGCCCCTGAAGCTGGCAGTTTCAGTCCCTTATTCTGCCCTGTGTG
GAAGCCCTGCTGAACCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCTTGAAA
TGCTGTGAGCTTGACTTGACTCCCAACCTTACCATGCTCCATCATACTCAGGTCTCCCTACT
CCTGCCCTGAGATTCTCAATAAGATGCTGTAACTAGCATTTTTGAATGCCTCTCCCTCCGC
ATCTCATCTTTCTCTTTTCAATCAGGCTTTTCCAAAGGGTTGTATACAGACTCTGTGCACTA
TTCTCACTTGATATTCAATCCCAATTCACTGCAAGGAGACCTCTACTGTACCCGTTTACTCT
TTCTACCTGACATCCAGAAACAATGGCCCTCCAGTGCACTACTCTTCACTTTGCTTTATG
GCCTTTCATCATAGTTGCCCATCCCTCTCCTTACTTAGCTTCCAGGTCTTAACTTCTCTG
ACTACTCTGTCTTCTCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCACTGCA
TCCATTTGTAGATTTTGTCTTCTCTCAGTTTACTCATTTGCCCTGGAACAATCACTGACA
TCTACAACCATACCATCTCACTAAATAAGACTTTCTATCCAATAATGATTGATACCTCAAA
TGTAATAAA

FIGURE 94

MGLQACLLGLFALILSGKCSYSPEDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLS
ELVQAVSDPSSPQYGKYLTLLENVADLVRPSPLTLHTVQKWL LAAGA QKCHSVITQDFLTCWL
SIRQABLLLPGAEFHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGG LHRFPPTSSLRQRP
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVSGGTSNNSQACAQFLEQYFHDSDLAQFMRLFG
GNFAHQASVARVVGQQGRGRAGIEASLDVQYLSAGANISTWVYSSPGRHEGQEPFLQWLML
LSNESALPHVHTVSYGDDDESLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH
QFRPTFPASSPYVTTVGGSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP
HLPPSSYFNASGRAYPDVAALSDGYVWVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSG
RPPLGLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGFGWDPVTGWGTPTSQLC

Signal sequence:

amino acids 1-16

N-glycosylation site.

amino acids 210-214, 222-226, 286-290, 313-317, 443-447

Glycosaminoglycan attachment site.

amino acids 361-365, 408-412, 538-542

Casein kinase II phosphorylation site.

amino acids 212-216, 324-328, 392-396, 420-424, 525-529

N-myristoylation site.

amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225,
248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488,
521-527, 533-539, 549-555

THE UNIVERSITY OF CHICAGO

FIGURE 96

MAGIPGLLFLFFLLCAVGQVSPYSAPWKPTWPAYRLFVVLPQSTLNLAKPDPFGAEAKLEVS
SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIYILSSSGDGAQHRDSGSSGKS
RRKRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKG
TQKLRVGFLKPKFKDGGRGANDSTSAMPEQMKEQWIRVKRTHVPKGWIKGNANDIGMDYDYA
LLELKPKPKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFGDVKDETYDLLYQQCD
AQPGASGSGVYVRMWKRQQQKWERKIIGIFSGHQWVDMNGSPQDFNVAVRITPLKYAQICYW
IKGNYLDCREG

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 93-97, 207-211

Glycosaminoglycan attachment site.

amino acids 109-113, 316-320

Casein kinase II phosphorylation site.

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

N-myristoylation site.

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

Serine proteases, trypsin family, histidine active site.

amino acids 171-177

FIGURE 97

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCCTGCTCCCCCGCCCCACCAGCC**ATG**GTGGTTT
CTGGAGCGCCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTGCTG
GCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTTCCCCAGCCTGTGGGAAGCCCCA
GCAGCTGAACCGGGTTGTGGGCGGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA
GCATCCAGAAGAAATGGGACCCACCCTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC
ACTGCTGCCCACTGTTTCAAGGACAACCTGAACAAACCATACTGTTTCTCTGTGCTGCTGGG
GGCCTGGCAGCTGGGGAACCTGGCTCTCGGTCCCAGAAAGTGGGTGTTGCCCTGGGTGGAGC
CCCACCCCTGTGTATTCTTGAAGGAAGGTGCCCTGTGCAGACATTGCCCTGGTGCCTCTCGAG
CGCTCCATACAGTTCTCAGAGCGGGTCTGCCCCTCTGCCTACCTGATGCCCTCTATCCACCT
CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCCTTGC
CCCACCCCTCAGACCCCTGCAGAAGCTGAAGGTTCCCTATCATCGACTCGGAAGTCTGCAGCCAT
TGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT
GGAGGGGGAGCGGGATGCTTGTCTGGGCGACTCCGGGGGCCCCCTCATGTGCCAGGTGGACG
GCGCTGGCTGCTGGCCGGCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCC
GGGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA
GCTCCGCGGGCGCGCTCAGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCCG
CCGCGCGCTCT**AG**GGCGCAGCGGGACGCGGGGCTCGGATCTGAAAGCGGGCCAGATCCACA
TCTGGATCTGGATCTGCGGCGGCCCTCGGGCGGTTTCCCCCGCCGTAATAGGCTCATCTACC
TCTACCTCTGGGGGCCCGGACGGCTGCTGCGGAAAGGAAACCCCTCCCCGACCCGCCCGAC
GGCCTCAGGCCCCCCCTCAAGGCATCAGGCCCGGCCAACGGCCTCATGTCCC GCCCCCCAC
GACTTCCGCCCCCGCCCCGGCCCCAGCGCTTTTGTGTATATAAATGTTAATGATTTTTAT
AGGTATTGTAAACCTGCCCACTATCTATTTATTCCTCCAATTTCAATAAATTATTTATT
CTCCAAAAA

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318
><subunit 1 of 1, 317 aa, 1 stop
><MW: 33732, pI: 7.90, NX(S/T): 1
M V V S G A P P A L G G G C L G T F T S L L L L A S T A I L N A A R I P V P P A C G K P Q Q L N R V V G G E D S T D S E W P
W I V S I Q K N G T H H C A G S L L T S R W V I T A A H C F K D N L N K P Y L F S V L L G A W Q L G N P G S R S Q K V G V A
W V E P H P V V S W K E G A C A D I A L V R L E R S I Q F S E R V L P I C L P D A S I H L P P N T H C W I S G W G S I Q D G
V P L P H P Q T L Q K L K V P I I D S E V C S H L Y W R G A G Q G P I T E D M L C A G Y L E G E R D A C L G D S G G P L M C
Q V D G A W L L A G I I S W G E G C A E R N R P G V Y I S L S A H R S W V E K I V Q G V Q L R G R A Q G G G A L R A P S Q G
S G A A A R S

Signal sequence:

amino acids 1-32

N-glycosylation site.

amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459,
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,
1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,
612-618, 623-629, 714-720, 873-879

FIGURE 99

GACGGCTGGCCACCATGCACGGCTCCTGCAGTTTCTTGATGCTTCTGCTGCCGCTACTGCTA
CTGCTGGTGGCCACCACAGGCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGATGGT
GGAGCTGCACAACCTCTACCGGGCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGA
GATGGGACGAGGAGCTGGCGCCTTCGCCAAGGCCTACGCACGGCAGTGCCTGTGGGGCCAC
AACAAGGAGCGCGGGCGCCGCGCGGAGAAATCTGTTTCGCATCACAGACGAGGGCATGACGT
GCCGCTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT
GCAGCCCAGGCCAGATGTGCGGCCACTACACGAGGTGGTATGGGCCAAGACAGAGAGGATC
GGCTGTGGTTCCCACTTCTGTGAGAAAGCTCCAGGGTGTGAGGAGACCAACATCGAATTACT
GGTGTGCAACTATGAGCCTCCGGGGAAACGTGAAGGGGAAACGGCCCTACAGGAGGGGACTC
CGTGTCTCCAATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGC
CCGAAGATGCTCAGGATTTGCCTTACCTGGTAACTGAGGCCCCATCCTTCGGGCGACTGA
AGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT
TGGTAAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGGAAAACCCAGGCC
CCAACCTTCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCGTAAC
AACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTA
CCTTCCCCAAATCGACCCATGTTCCTATCCCAAATCAGCAGACAAAGTGACAGACAAAACA
AAAGTGCCCTCTAGGAGCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG
GGAACCTCTACCCATGCCCAGGAGGAGGTGAGGCTGAGGCTGAGTTGCCCTCTTCCAGTG
AGGTCTTGGCCTCAGTTTTTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC
CACACGGGGCACACCTCCTCCAAGTCCCTGCCCAATTTCCCAATACCTCTGCCACCGCTAA
TGCCACGGGTGGGCGTGCCTGGCTCTGCAGTCGTCTTGCAGGTGCAGAGGGCCCTGACA
AGCCTAGCGTTGTGTACAGGGCTGAACTCGGGCCCTGGTCATGTGTGGGGCCCTCTCTGGGA
CTACTGCTCCTGCCTCCTCTGGTGTGGCTGGAATCTTCTGAATGGGATACCACTCAAAGGG
TGAAGAGGTCAGCTGTCTCCTGTCTATCTTCCCCACCTGTCCCCAGCCCCATAACAAGATA
CTTCTTGGTTAAGGCCCTCCGAAGGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCC
ATCTTGAGGCACAAGGCCCTGGCTGGCTGCGAGCTCAGGAGGCCGCTGAGGACTGCACACC
GGGCCCCACCTCTCCTGCCCTCCCTCCTGAGTCTTGGGGGTGGGAGGATTGAGGGAGCT
CACTGCCCTACCTGGCTGGGGCTGTCTGCCACACAGCATGTGCGCTCTCCCTGAGTGCCTG
TGTAAGTGGGGATGGGGATTCTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTC
TTTGAGTGGGGGAGGCAGGGACGAGGGAAGAAAGTAACTCCTGACTCTCCAATAAAAACTT
GTCCAACCTGTGAAA

FIGURE 100

MHGSCSFLMLLLPLLLLLLVATTGPGVGTDEEKRLMVELHNLVYRAQVSPTASDMLHMRWDEE
LAFAKAYARQCVWGHNKERRRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQ
MCGHYTQVVWAKTERIGCGSHFCEKLGVEETNIELLVVCNYPEPPGNVKGKRPYQEGTPCSQC
PSGYHCKNSLCEPIGSPEDAQDLFYLVTPEAFSFRATEASDSRKMGTTPSSLATGIPAFVLVTEV
SGSLATKALPAVETQAPTSLATKDPSPMATEAPPVCTTEVPSILAAHSLPSLDEEPPVTFPKS
THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEALPPSSEVLAS
VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSV
SGLNSGFGHVWGPLLGLLLLPLVLGIF

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 114-118, 403-407, 409-413

Glycosaminoglycan attachment site.

amino acids 439-443

Casein kinase II phosphorylation site.

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

N-myristoylation site.

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237,
250-256

Amidation site.

amino acids 82-86, 172-176

Peroxidases proximal heme-ligand signature.

amino acids 287-298

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.

amino acids 127-138

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.

amino acids 160-172

FIGURE 101

GTAACCTGAAGTCAGGCTTTTCATTTGGGAAGCCCCCTCAACAGAAATTCGGTCATTCCTCAAGT**TAT**GTGGACGT
 ACTTCTGTGTCTCCCTCTGCTGTCTTTTTCACATTAGCAGACCGGACTTAAGTCACAAACAGATTATCTTTCAT
 CAAGGCAAGTTCATGAGCCACCTTCAAAGCCTTCGAGAAGTGAACCTGAACACAAATGAATTGGAGACCAATCC
 AAATCTGGGACAGCTCTCGGCAAAATATTACACTTCTCTCTTGGCTGGAAACAGAGATTGTTGAAATACCTCCCTGA
 ACATCTGAAGAAGTTTCAGTCCCTTTGAAACTTTGGACCTTAGCAGCAACAATATTTTCAGAGCTCCAACTCGAATT
 TCCAGCCCTTACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAAACCTGGGTATTTTGAACA
 TTTGGCCAAACAGATCTCTGTGTGTTAAAGCTGAACAGGAACGAAATCTCAGCTATCCACCACCAAGATGTTTAAACT
 GCCCAACTGCAACATCTCGAATTGAACCGAAACAGATTAAAATGTAGATGGACTGCATCTTCCAGGCTCTTGG
 TGCTCTGAAGTCTCTCGAATTGCAAGAATAAGGAGTAACTGAGTAACTTATGGATGGAGCTTTTGGGGCTGAGCAA
 CATGGAAATTTTGGAGCTGGAGCAATAACAACCTTAACAGAGATTACCAAGGCTGGCTTTACGGCTTCTGTGTGCT
 GCAGAGAACTTCACTCGCCAAAATGCATCAACAGGATCAGCCCTGATGCTGGAGTCTCTGCCAAGAGTGCAG
 TGAGCTGGACCTAACTTCAATCACTTATCAAGGTTAGATGATTCAGGCTTCTTGGGCTTAAGCTTACTAAATAC
 ACTGCACATTGGGAACAACAGAGTCAGCTACATTTGCTGATTGTGCTTCCGGGGCTTCCAGTTTAAAGACTTT
 GGATCTGAAGAACAATGAAATTTCTCGGACTATTGAAGACATGAATGGTGCTTCTCTGGGCTTGACAAACTGAG
 GCGACTGATACTCCAAGGAAATCGGATCCGTCTATTACTAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCA
 TCTAGACCTGAGTGACACGCAATCATGTCTTTTACAAGGCAATGCATTTTACAAATGAAGAACTGCAACAATT
 GCATTTAAATACATCAAGCTTTTGTGCGATTGCCAGCTAAAATGCTCCACAGCTGGGTGGCGGAAACCAACT
 TCGAGAGCTTTGTAATGCACTTTGTGCGCAATCTCAGCTGCTTAAAGGGAAGAGCAATTTTGTGTTAGCCCAGA
 TGGCTTTGTGTGTGATGATTTTCCCAAACCCAGATCAGCGTTCAAGCCAGAAACAGCTCGGCAATAAAGGTTCC
 CAATTTGAGTTTTCATCTGCTCAGCTGCCAGCAGCAGTGATTTCCCAATGACTTTTGTCTGGAAAAAGACAATGA
 ACTACTGATGATGCTGAAATGAAAAATATGACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCA
 CATCTCTCGGCTGCGGAGGTGGAAATTTGCCAGTGAGGGGAAATATCAGTGTGTCATCTCCAATCACTTTGGTTT
 ATCTACTACTCTGAAGCCAGCTTTCAGTAAATATGCTTCCCTCAATCACCAGACCCCAATGGATCTCAGCAT
 CCGAGCTGGGGCCATGGACCGCTTGGAGTGTGCTGCTGTGGGGCACCCAGCCCCAGATAGCCTGGCAGAGGA
 TGGGGGCAAGCACTTTCCAGCTGCAAGGGAGAGACGATGCTGTGATGCGCGAGGATGACGTGTTCTTTATCGT
 GGAATGGAAGATAGAGGACATTTGGGTATACAGCTGCAAGCTCAGAACAGTGCAGGAAGTATTTTCAAGCAATGC
 AACTCTGACTCTCTAGAAACACCATCATTTTTCGGGCCACTGTGTGGACCGAAGCTTAACCAAGGGGAGAAACAGC
 CGTCTTCAAGTGCATTGCTGGAGGAAGCCCTCCCTTAACTGAAGTGGACCAAGATGATAGCCCATTTGTGGT
 AACGAGAGGCACTTTTTCGAGCAGGCAATCAGCTTCTGATTAATTGTGGAATCAGATGTCAGTGTGCTGGGA
 ATACACATGTGAGATGCTAAACACCTTGGCACTGAGAGAGGAAACGTCGCGCTCAGTGTGATCCCACTCAAC
 CTGCGACTCCCTCAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGCTGTGATCATAGC
 CGTGGTTGTGTGTGTGGGCAGTCACTCGTGTGGTGGTCACTATACCAACAAGGCGGAGGAATGAAGA
 TTGCGACATTTACCAACAGAGATGAGACCAACTTGCAGCAGATATTCTCAGTATTTTGTCACTCTCAGGGAACGTT
 AGCTGACAGGCGAGATGGGTACGTCGCTTCAGAAAGTGGAGCCACCAACCTTTGTCACTTTGAGTGTGCTGG
 ATTTTCTTACCACAACATGACAGTAGTGGGACCTGCCATATTGACAAATAGCAGTGAAGCTGATGTGGAAGCTGC
 CAGAGATCTGTTCTTTGTGCTTTTGGGATCCACAGGCCCTATGTGATTTTGAAGGAAATGTGTATGGCTGCA
 TCCTTTTGAACAATATCATACAGTGTGAGTCTGACCTGACCCAAAGAACAGTTTAAATGGACCACTATGAGCCGAGTTA
 CATAAAGAAAAAGGATGCTACCGATGTTCTCATCTTCAAGAAGATCCAGACGGAGCTCAGTAATATATCT
 GTGGCCTTCAATGTGAGGAAGCTACTTAACTAGTTACTCTCAATGAAGGACCTGGAATGAAAAATCTGTG
 TCTAAACAGTCTCTTTGATTTTATGTGCAAAATCCAGGCCAGCTCGGTGGCTCGAGTAATCTTTTATGGG
 TACCTTTGAAAGCTCTCAGGAGACCTCACCTAGATGCCATTCAAGCTTTGGACAGCCATCAGATTGTGAGCC
 AAGAGCTCTTTTATTTGAAGCTCATTTCTCCACAGACTTTGGACTCTGGGTGAGGAAAGATGGAGAAAGGAG
 AGATTTTCAAGGAAGAAATTCACATTTGTACCTTTAAACAGACTTTAGAAAACCTACAGGACTCCAAATTTTTCAGT
 TTATGCTTGGACAT**TAG**ACTGATGAGACCAAGGAAAGCTTAACTACTACTCACTCAAGTGAACCTTTTATTTA
 AAGAGAGAGATCTTATGTTTTTAAATGGAATGATGAATTTTAAAGGATAAAATGCTTTTATATACAGAT
 GACCAAAATTTACAAAGTTATCAAAATTTTATATCTGGGAATGATGCTCATATAAGAAATACCTTTTAAACTA
 TTTTTTAACTTTGTTTATGCAAAAGTACTCTACGTAAATTAATGATATAAATCATGATATTTTATGATATT
 TTAAATCCGAGATTTCTTTTATGGAAATGAGTACTAAAGCATTTTAAATAATCACTGCCCTGTACCAATTTT
 TTAATAGAAAGTACTTCTATTATATTTTGCACATTATATTTAATAAATGTGTCAAATTTGAA

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

FIGURE 102

MVDVLLLFSLCLLPHISRDLSDHNRSLFSIKASSMSHLQSLREVKLNNNELETIPNLGPVSAN
ITLLSLAGNRIVEILPEHLKEFQSLDTLSSNNISELQTAFPALQLKYLNLNRRVTSMEP
GYFDNLANTLLVLKLNRRNRISAI PPKMFKLPQLQHLELNRRNKIKNVDDLTFQGLGALKSLKM
QRNGVTKLMDGAWGLSNMEILQLDHNNTETITKGWLYGLLMLQELHLSQNAINRISPDWE
FCQKLSLELDLTFNHLRLDDSSFLGLSLNLTLLHIGNNRVSYIADCAFRGLSSSLKTLDLKNNE
ISWTIEDMNGAFSGLDKLRRLILQGNRIRSIITKAFTGLDALEHLDLSDNAIMSLQGNAFSQ
MKKLQQLHLNTSSLLCQCQLKWLPPQWVAENNFQSFVNASCAPQLLKGRSIFAVSPDGFVCD
DFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKNELLHDAEMENYAHLRAGQ
GEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLITRAGA
MARLECAAVGHPAPQIAWQKDGDTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQN
SAGSISANATLTVLETPSFLRPLLDRTVTGKETAFLQCIAGGSPPPKLNWTKDDSPLVVTER
HFFAAGNQLLIIVDSVSDAGKYTCMSNTLGTERRGNVRLSVIPTPTCDSPQMTAPSLDDDG
WATVGWVVIIVVCCVVGTSLVWVVIIYHTRRRNEDCSITNTDETNPADIPSYLSSQGTAD
RQDGYVSSSESGHHQFVTSSGAGFFLPQHDSSGTCHIDNSSADVEAATDLFLCPFLGSTGP
MYLKGNVYGSDFPFTYHTGCSFDPRTVLMDHYEPSYIKKKECYPCSHPEESCESRFSNLSW
PSHVRKLLNTSYSHNEGPGMKNLCLNKSLLDFSANPEPASVASSNSFMGTGFKALRRPHLDA
YSSFGQPSCQPRAFYLLKAHSSPDLDGSGSEEDGKERTDFQENHICTFKQTLNRYRTPNFQS
YDLDT

Signal sequence:

amino acids 1-19

Transmembrane domain:

amino acids 746-765

N-glycosylation site.

amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459,
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,
1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,
612-618, 623-629, 714-720, 873-879

THE UNIVERSITY OF CHICAGO

GGGGAGGAGGATTTGACCACTAAAGAGGAGACTTTTTTTGGTGTGGTGTGGCTTGTGGTGTCTTGCACAAATATG
 AAGGATTCAGGACGCGACCTTTCTCTGGACCGAAGCCGATCGTAACATCTGATGTGCACAGAGAGAAGCAAGAC
 GAAGCTTTTCTTGTAGCGCGCTGGATCTTAACAACAAATGTCTATATGTGCACACAGGGAGCATTCAAGAAATGAAC
 CATTACCAAGTGTAGACGCGCGGGGTGGTGTGTTCTTGACATATAAATAATCTTAAGACAGCTGTCTCCCTCG
 CCAACGGAAGAAAAAGGATGATGGAAATGAAGAACCGGAGGATCTCAAGAAAAAGGATTTGTTCATTTTCTC
 TATAAGGAGAAGATGTAGCCCAAGGATATTTTGGAAATGAAAGTTGTGGGGCTTTTGTAGTAAGATGAAGAAGCT
 GGTGTGGTGGTGTCTTTCTTTCTTTCTTTGAAATTTCCACAGAGGAGAGAGAAATAATAATCATCTGCAGAAATCT
 TTTACAGAGAAGAAATGTACCGCGCGAGCAATGTAGGACGATGTATGGGGGAGAGAACCAACGAGAGCAGTGTGGA
 TTTGTGGCTATTTGTACTAAATTCAGCGAATTCAGTGTGGATTTTCTCTCAATCAACCTCTTTTTTTAAAT
 TTTTATTTCTTTGGTATCAAGATCATGCGTTTCTCTGTGTTCTTAACCACTGGATTTCCATCTGGATGTGTGCT
 GTGATCAGTCTGAAATAACAATGTTTGATTTCCAGAGAGGACCAACAGAGATAAATATGAGATTTTGAACAGAT
 GAGCTCTACATCCACAGCAGATGATGATGTGCTAGGTTTAAACGGGCCCTATTGACCGCCTGCTGTGGTGTCT
 GCTGGCTTCTCAACTCTTGTGGTGGTGGTGTGGTGTGGCGGCTCAGACTGCCCTTCTGTGTGCTCTCGACGAA
 CCGATTCAAGCAGGTGATTTGTGTTCGGAAAAACCTGCGTGAGAGTCCGGATGGCATCTCAACCAACACCGCT
 GTGAACCTTCATGAACTTAATCCAGATCATCAAGGTGAACAGCTCAAGCACTTGAAGCATTTGAAATCTCT
 ACAGTTTGTAGTAGGACCATATACAGAACCATTTGAATTTGGGGCTTCAAGTTGTCGGCAACTCAACATCTGGA
 ACTCTTTGACATCGTCTTACTACATCCGGAATGGAGCTTTGTATCTGTCTAAACTGAAGGAGCTGTGTT
 GCGAAACACCCCATTTGAAGAGCATCCTCTTATCTCTTATACAGAAATCTCTTTCTTGGCCGACAGACTAGACTTAGG
 GGAATTTGAAAGACTTTTATACATCTCAGAAGTGCCTTTGAAGTCTTCCCACTTGGATTTGAACTTGT
 CATGTGCAACTCTCGGGAATCCCTAACCTCACACCGCTCATATAAACTAGATGAGCTGGATCTTCTGGGAATCA
 TTATCTGCCATCAGCGCTGGCTCTTTCTCAGGGTTGTGATCGACTCAAAATCTGGATGTATCAGTCCAGAT
 TCAAGTGTGAACGAGATGCCCTTTGACAACCTTCAGTCACTAGTGGAGATCAACTGGCACACAATAATCTAAC
 ATTACTGCTCATGAACCTCTCACTCTCTGTCATCATCTAGAGCGGATACATTTACATCAACCTTGGAACTG
 TAACCTGTGACATCTGTGCTCAGCTGGGTGATATAAGAGCATGGCCCTCGAACACAGCTTTGTGTGGCGGGT
 TAAACACTCTCCAATCTAAAGGGGGAGGTACATTTGAGAGCTGACCAGGAATACTTCACTGATCATGTCTCGGT
 GATTTGGAGCGCCCTCGAGACCTCAATGTCTCACTGAAGGCACTGGAGCTGAGGCTGAAATGTGGGGCTCTCAATC
 CTGTCATCTGTATCTTGATTAATCTCAATGGAACAGTCATGACATGGGCGGATCCAAAGTGGCGGATAGCTGT
 GCTCAGTGAATGGTACGTATAATTTCACAAAATGTAACGTGACAGATCAGGCACTGTACATGATGTGTGAGTAT
 TTTCCGTTGGGAATACTACTGTTCTCAGCCACCTGAAATGTTACTGACGACCAACTACTCTTTCTCTTATTTT
 AACCGTCCAGTGAAGACATTTGAACCGTCTCAGGATGAGGACCGGACACAGATACAATGTGGGCTCCACTCC
 AGTGGTCAGCTGGGAGACACCAATCTGTGACCCACTCTCTCACACACAGAGCAAGGTGCAGACAGAAAACTT
 CACCATCTCAGTGAATGATAAATACAGTGGGATCCAGGAATGTATGAGCTCATGAAGATCCCAAAATCATCAT
 TGGGTTGTTTGGGCCATACACTCATGCTCGAGTGTCTGTCATTTTCAACAGATGAGGAAGCAGCACCA
 CTGGCAAAACACTCAGCCCCACAGGACTGTTGAAATTTAATGTTGGATGTAGATTCAGGGAGACACACC
 CATGGAAGGCCACTCGCATGCTCTCTCATCGACTGAGCAGCTAAATCACTATACTCATACAATCTCTCTT
 CAACCAACCAACAGCTTTACCAATAAATTAATCAATCAGCTTCAGTGCATGAACCAATTTATGTATCGGAATGA
 CTCTAAAGCAATGTACAGAGACTCAAAATTAAGCAATTTACAGAGTTACAAAAACCTTAATCAAAAGAAA
 GACAGTTTATTAATAATGACAAATGACTGGCTTAATCTACTGTTTCAAAAAGTGTCTTACAAAAAACCA
 AAAAAAGAAAGAAATTTATTATTAATAAATCTGATTTGATCTAAGACAGACAAAA

FIGURE 104

MLNKM TLHPQQIMIGPRFNRALFDPLLVL LALQLLVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGI STNTRLNLHENQIQI I KVNSFKHLRHL EILQLSRNHRTIETIGAFNGLA
NLNTLELFDNRLLTTPNGAFVYLSK LKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNLRYLNLAMCNLR EIPNLTPLIKLDEL DLSGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHPWNCNC DIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRAS TSLTSVSWITPNGTVMTHGAYKVRIAVLS DGTLNFTNVTVQD TGMVTCMVSNSVGN
TTASATLNVTAATTPFSYFSTVTVETMEPSQDEARTDNNVGPTPVVDWETTNTVTTSLTPQ
STRSTEKTFTIPVTDINSGIPGIDEVMKTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYSYKSPFNHTTNTVNTINSIHSS
VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438,
442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,
391-397, 422-428, 433-439, 531-537

06-07-2008

AGCGAGCGCTGCTCAAGGCGCAACTCTGTGTGCAGTTGGCAGTTCCTTTTCGGTTTCCTCTCGTGTTTGGGGGGA
TAAAGGGCTTCGCGCGCGGAGTAAAGAGGAAATGACGCGGAGCGAGCGAGGAGGAGCGCGACGCGACCG
GAGCGCGCGCGTGCAACCTCGGCTGCGAAGTTTGTCCGCGCCGAGCGCGCGCGCTGGAGACTTCGGGTGA
GACTAGCGCGCTGACGACCGCGATGAGCGCGCGCGAGCTCCGTGTCCGCGCGCGCGGGTGGGGCTGCTGTGTGC
CGGTGTCTGGGGCGCGTGGCGCGTCCGACGACGCGCGTCCGGGAGACTCGGGGAGCCCTCTGGGGTAGCGCCG
GAGCGCCGATCGCCCACTACGCTGCGCTCGCTCGGGGACCTCTGTGACATCGAGCTCGTAAACGCGCTGACGCGTCTT
CCCGAGCGACTCCCGCTCTGGGTGCTCGGTGTGACTTAAGTCACAACAGATTTCTTCAACAAGSCAGACTCC
ATGAGCCAGCTCTCAAGGCTTCGAGAGATGAAATCGAACCAATGAATTAAGGACACCTTCAAAATCTGGACCA
GTCTCGGCAAAATATACITTCCTCTCTGTGGTGGAAACAGAGTTTGAATACTCCCTGAACATCTGAAGA
TTTCAGTCCCTTGAACATTTGGAATCTTAGCAGCAACAAATTTAGAGCTCCAAATGCAATTTCCAGCGCTTACA
CTCAAATATCTGTATCTCAACGCAACCGAGTCACATCAATGGAACCTGGGTATTTGCAAAATTTGGCCAAACA
CTCTCTGTGTAAAGCTGAAACGAGAACCGAATCTCAGATCTACCCAACCAAGATTTTAAATCGGCCCACTGCAA
CATCTCGAATGAGCGCAACAGATTAATAATATGATGAGTACATTCAGAGGCTTGGTGTCTCTGAGTCTCT
CTGAAATTCGAAAGAAATGGAGTATGCAACAACTATGATGTGACCTTTTGGGGCTGAGCAACATGGAAATTTTG
CAGCTGGACATCAACCACTAACACAGATTAACAAAGCTGGCTTTTACCGCTTCTGATGCTGACGAAATTCAT
CTCAGCAAAATGGCCATCAACAGGATCAGCCCTGATGTCTGGGAGTCTGCTCAGAGAACTCAGTGAAGCTGACCTA
ACTTTCAATCAACTATCAAGGTTAGATGATTCAGCTCTCTGGCTTACGCTTAATAATCACTGCAATGGG
AACACAGAGATCGACTACATGCTGATTTGGCTCTCGGGGGCTTCCAGTTTAAAGACTTTGGATCTGATGAAC
AATGAATTTCTTGCACTATTAAGAACATGAATGGTGTCTCTCTGGGCTGACAACCTGAGCGAGCTGATGACT
CAAGGAAATTCGATCCGTTCTATTAATAAAGATGCTCATCTGGTTTGTGATGATTTGGAGACTCTAGACCTGAGT
GACACGCAATCATGTCTTTCAAGGCAATGATTTTCAAAATGAAGAACTGCAACATTCGATTTAAATACA
CTAAGCTCTTTGTGCGATGCGCACTAAATGCTCCACAGTGGGTGGCGAAAACACTTTCAGAGCTTTGTA
AATGCGAGTGTGCCATCTCAGCTGTCTAAAGGAGAAAGCACTTTTCTGTAGGCCAGTGGCTTGTGTGTT
GATGATTTTCCAAACCCAGATCAGCGTTGAGCGAAGAACAGTGGCGCAATAAAGGTTTCCAAATTTGATGTTT
ATCTGTCTCAGCTGCGGACCGAGTGAATTTCCCAATGACTTTGTCTGAAAAGAACAAATGAACATCTGCAATG
GCTGAATGGAAAAATTTGACACACTCGCGGCCAGGCTCAAGTGGCGAGGATGATGAGATTAACCAACTCTTGGCTG
CGGAGGTGGAATTTGGCAGTGAAGGGAATAATCAGTGTGTCTATCCCAATCTTTGGTGTCTATCTCTGTC
AAAGCGAAGCTTACAGTAATATGCTCTCCCTCATCAACAGAACCCCATGATCTACCACTCGAGCTCGGGCC
ATGCGACCTGGAGTGTCTGTCTGGGGGACCCAGCGCCCGAGATAGCTTGGCAAGGATGGGAGCGACGAG
TTCGAGCTGCGACGGGAGAGCGCATGTGATGCTCCGAGGATACGCTGTTCTTATCTTGGGTGTGAAGATA
GAGGACATTTGGGTATACAGCTGACAGCTGACAGCACTGACGAGCAATTTTCAGCAAACTGCACTGTCTC
CTAGAACACCATATTTTGTGGGCGCATGTGGACCGAATCTGAACACAGGAGAAACAGCGCCTCTCACTAGTGC
ATTGCTGGAGGAAGCTTCCCTTAACTGAACTGACCAAGATGATAGCCCATTTGGTGTGTAACCGAGGAC
TTTGTTCAGCGAGCAATCAGTCTCTGATTTTGTGGACTCAGATGTGATGATGCTGGGAAATACAGTGTGAG
ATGTTCTAACCACTTGGCATGAGAGAGAAACGTGGCTCAGTGTGATCCCACTCAGCTCAGCTCCCTT
CAGATGACAGCCCATCTGTGATAGACGATCGAGTGGGCGATTTGGGCTGTGATCTCAGCTGCTTTTGAACAC
GTGTGGGCAAGCTCAGTGTGTGGGTGATCATATACCAACAGCGGAGGATGAAGATTCGACGATTACC
AACACAGATGAGCAACATTCGACAGATATTCCTAGTTATTGTTCATCTCTGCGGAAGCTTAGCTGACAGGCG
GAGTGTGCTGTGTTCTGAAGATGGAGAACCAACAGCTTTGACATCTCAGTGTGGAGCTGACAGACTGTCT
CAACCATGACAGTATGGAGCTTCCCATATGCAATAGCAGTGAAGCTGATGTGGAGCTGAGATTTCTTACCA
CTTTGTTCGTTTGGGATCCACAGCCCTATGATTTGAAGGGAATGTGTAATGCTCAGATCTTTTGAACAC
TATCATACAGGTGAGCTCTGACCCAGAGACGATTTTAATGGAACCATATGAGCGGCACTACATAAAGAAAAAG
GAGTGTACCACTTCTCATCTCTCAGAGAACTCTGCGACAGGAGCTTCAGTATATATCTGTGGCTTTCACAT
CTGAGGAAGCTACTTAAACACTAGTATCTTCAAAATGAAGGACCTGGAATGAAAAATCTGCTGTCTAAACAGTCC
TCTTAGATTTTAGTGCACAACTCCAGCGCGAGCTGGTGGTCTCGCTCGAGTAATTTCTTCATGGGTACTTTGAAA
GCTCTCAGAGACTCAACTAGATGCTTATCAAGCTTTGGACAGCATCAGATTTGACGCAAGAGCTTTTAT
TGAAGCTCTCATCTTCCCGAGACTGGACCTGGGTGAGGAAAGTGGGAAAGGAGACAGATTTTCAGGA
GAAAATCACTTTGTATCTTTAAAGCACTTTAGAAAATCAGGACTCGAAATTTTCACTCTATAGCTTGGAC
ACATATGCTGATGAGCAAGGAAAGACTTAACTACTACTCAGTGAATGACITTTATTTAAAGAGAGAGAT
CTTAGTGTTTTAAATGGAGTTATGAATTTTAAAGAGGATAAATACGTTATTTATACAGAGCAAAATATAC
AAAAAGTATGAAATTTTATATCGGGAATGAGTCTCATATAGAATACCTTTTAAACTATTTTAACTCTT
TTTATGCAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATATTTATGTATTTTATATGCGAGA
TTCTTTTATTGGAATAGATGATCAAGACATTTTAAATATACCTGCTTTACCATTTTAAATAGAGATTT
ACTTCATATATTTTGCACTATATTTTATAAATGTGCAATTTGAAAAAATAAATAAATAAATAAATAAATAA

FIGURE 106

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGRGELQPSGVAAERPCPTTCRCLGDLDDCSR
KRLARLPEPLPSWVARLDLSHNRLSFIKASSMSHLQSLREVKLNNNLEETIPNLGPFVSANIT
LLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGY
FDNLANTLLVLKLNRRIRISAIPPKMFKLPLQLQHLELNRNKIKNVDGLTFQGLGALKSLKMQR
NGVTKLMDGAFWGLSNMEILQLDHNNTETKGLWLYGLLMLQELHLSQLNAINRISPDWFEFC
QKLSELDLTFNHL SRLDDSSFLGLSLNTLHIGNNRVSYIADCAFRGLSSSLKTLDLKNNEIS
WTIEDMNGAFSGGLDKLRLILQGNRIRISITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQMK
KLQQLHLNTSSLLCDCQLKWL PQWVAENNFQSFVNASCAPQLLKGSRIFAVSPDGFVCDFF
PKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKNELLHDAEMENYAHLRAGGE
VMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVAKLTVNNMLPSFTKTPMDLTI RAGAMA
RLECAAVGHPAPQIAWQKDGTDFFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQNSA
GSISANATLTVLETPSFRLPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTERHF
FAAGNQLLIIVSDVSDAGKYTCESNTLGTGERGNVRLSVIPTPTCDSPQMTAPSLDDDGWA
TVGVVIIAVVCCVGTSLVWVVIIYHTRRRNEDCSITNTDETNLPA DIPSYLSSQGT LADRQ
DGYVSSSESGHHQFVTSSGAGFFLPQHDSSTCHIDNSSSEADVEAATDLFLCPFLGSTGPMY
LKGNVYGSDDPFETYHTGCSFDPRTVLMDHYEPSYIKKKECPCSHPSESCERSFSNISWPS
HVRKLLNTSYSHNEGPGMKNLCLNKSSLD FSANPEPASVASSNSFMGTFGKALRRPHLDAYS
SPGQPSDCQPRAFYLLKAHSSPDLDGSGSEEDGKERTDFQEENHICTFKQTLNRYRTPNFQSYDLDT

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 808-828

N-glycosylation site.

amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519,
688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

Glycosaminoglycan attachment site.

amino acids 886-890

Casein kinase II phosphorylation site.

amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378,
383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735,
799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022,
1073-1077, 1079-1083, 1081-1085

Tyrosine kinase phosphorylation site.

amino acids 667-675

N-myristoylation site.

amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433,
513-519, 588-594, 672-678, 683-687, 774-780, 933-939

Leucine zipper pattern.

amino acids 58-80, 65-87

FIGURE 108

MEGEEAEQPAWFHQFWRPGASDSAPPAGTMAQSRVLLLLLLPPQLHLGPVLAVRAPGFGRS
GGHSLSPREENEFAEEEPVLVLSPEEPGPGPAAVSCPRDCACSQEGVVD CGGIDLREFP GDLP
EHTNHL SLQNNQLEKIYPEELSR LHRLET LNLQNNRLTSRGLPEKA FEHLTNLNYLYLANNK
LTLAPRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDNMFGSSNV
EVLILSSNFLRHVPKHLPPALYKHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDN
ETFWKLSSELYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLHSN
QLREQGIHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMILHNQITGIGREDFATTYF
LEELNLSYNRITS PQVHRDAFRKLRLLRSLDLSGNRLHTLPPGLPRNVHLKVKRNELAALA
RGALAGMAQLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYL
QNNKISAVPANAFDSTPNLKGIFLRFNKLAVG SVVDSAFRR LKHLQVLDIEGNLEFGDISKD
RGR LGKEKEEEEEEEEEEEETR

Signal sequence:

amino acids 1-48

N-glycosylation site.

amino acids 243-247, 310-314, 328-332, 439-443

Casein kinase II phosphorylation site.

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

N-myristoylation site.

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341,
477-483, 498-502, 539-545, 548-554

Leucine zipper pattern.

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493,
535-557

FIGURE 109

GGGAGGGGGCTCCGGCGCGCGCAGCAGACCTGCTCCGGCCGCGCGCTCGCCGTCTCTCCGGAGCGGCAG
CAGTAGCCGGGCGCGCAGGGGCTGGGGGTTCCTCGAGACTCTCAGAGGGGCGCCTCCCATCGGCGCCCCACACCC
CAACCTGTTCCTCGCGGCCACTGCGCTGCGCCCCAGGACCCGCTGCCCAACATGGATTTCCTCTGGCGCTGGT
GCTGGTATCTCGCTTACTCTCAGGCGCGCGCGCGAGTTCGACGGGAGGTGGCCAGGCAAAATAGTGTCTATCGAT
TGGCGTATGTCTTATGGTGGGAGGATTGACTGTCTGCTGGGGCTGGGCTCCCAAGCTCTTGGGGACAGTGTTCAGCC
TGTGTGCCAACACGATGCAAACTATGGTGAATGTATCGGGCCAAACAAAGTGCAAGTGTATCTCTGGTTATGTCTGG
AAAAACCTGTAAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCCCGCCCTGAAGACAGGGTGCAATGAACACTTA
CGGCAGCTCAAAAGTGTACTGTCTCAACGGATATATGCTCATGCGGGATGGTTCCTGCTCAAGTGCCTGACCTG
CTCCATGGCAAACTGTAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCTGGCCT
GCACCTGGCTCTGATGGGAGGACCTGTGTAGATGTGATGAATGTGCTACAGGAAGGCCCTCTCTGCCCTAGATT
TAGGCAATCTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTCAATAAGGCTTCGATCTCATGTATATTGAGG
CAAAATATCAATGTCAATGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAA
CGTACGTGGGTCTTACAAGTGCAAAATGTAAGAAGGATACCGGGTGATGGAATGACTGTTGTGTATATCCCAAA
AGTTATGATTGAACCTTCAGGTCCAATTCATGTACCAAAGGGAATGGTACCATTTTAAAGGGTGACACAGGAAA
TAATAATTGGATTCTGTATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCTCTATCATATTACCAA
CAGGCCACTTCTAAGCCAACAAACAGACCTACACCAAGGCCAACCAAAATTCCTACTCCACCACCAACCAACC
CCTGCCAACAGAGCTCAGAACACCTTACCACTACAAACCCAGAAAGGCCAACCCGGACTGACAACTATATAGC
ACCAGCTGCCAGTACACCTCCAGGAGGGGATTACAGTTGACAAACAGGGTACAGACAGACCTCAGAAACCCAGAG
AGATGTGTTTCACTGTTCTGGTACACAGTTGTAATTTTGACCATGGACTTTGTGGATGGATCAGGGAAGAAAGACA
TGACTTGCATGGGAACCAATCAGGAGCCAGCAGGTGGACAAATATCTGACAGTGTGCGCAGCCAAAGCCCCAGG
GGGAAGAACCTGCAGCTTTGGTGCTACCTCTCGGCCGCTCATGCTTCAGGGGACCTGTGCCCTGTCAATCCAGG
CAAGGTGACGGGGCTGCACCTTCGGCAGACTCCAGGTGTTTGTGAGAAAACACGTCGCCACGGAGCAGCTGGT
GGGAAGAAATGCTGGCCACTGGCTCGGAGGCAACACAGATCACCTTCGGAGGGGCTGACATCAAGAGCGAATCACA
AAGATGAATTAAGGGTTGGAaaaaaaAGATCAATGATGAAAAATTAAGGAATCTGGGATATTGAGCTCGAGAG
AAGAGACTGAGGGCAAAACATTGATGGTTTCAAGTATATGAAGGGTTGGCAACAGAGGGGTGGCGACCACTG
TTCTCCATATGCACTAAGATAGAACAAAGAGGAACTGGCTTAGACTAGAGTATAAGGGGATTTCTTGGCAGG
GGCCATTGTAGAACTTTCATAAAAAAAGAGGTGTAAAAATCTCAGTATCTCTCTCTTTCTAAAAAAATTAGA
TAAAAATTTGCTATTTTAAGATGGTTAAAGATGTTCTTACCCAGGAAAAAGTAAACAAATATAGAAATTTCCCAAA
AGAATGTTTGTATCTACTAGTAGTATGTCAGTGAAAAATCTTTAGAACTAAATTAATTTGGACAAGGCTTAATTTAGG
CAITTCCTCTTGACCTCTCAATGGAGAGGGATTGAAAGGGGAAGGCCCAAAATGCTGAGCTCACTGAAATA
TCTCTCCCTTATGGCAATCCTAGCAGTATTAAGAAAAAAGGAAACATATTTATCCAAATGAGAGTATGATGGAC
AGATATTTTAGTATCTCAGTAATGTCTAGTGTGGCGGTGGTTTCAATGTGTTCTCATGGTAAAGGTATAAGCC
TTTCATTGTGTTCAATGGATGATGTTTCAGATTTTTTTTTTTTTTAAAGAGATCCTTCAAGGAACACAGTTTCAGAGAG
ATTTTCATCGGGTGCACTCTCTGCTTCGTGTGTGACAAGTTATCTTGGCTGCTGAGAAAGAGTGCCCTGCCCC
ACACCGGCAGACCTTTCCCTCACCTCATCAGTATGATTCAGTTTCTCTTATCAATTTGAACTCTCCAGGTTCCAC
AGAACAGTAATATTTTTTGAACAATAGGTACAATAGAAAGTCTTCTGTCATTAAACCTGGTAAAGGACAGGCTGG
AGGGGGAAAAATAAATCATTAAGCCTTTGAGTAACGGCAGAATATATGGCTGTAGATCCATTTTTAATGGTTTCATT
TCCTTTATGGTCATATAACTGACAGCTGAAGATGAAAGGGGAAAAATAAATGAAGAAATTTTACTTTTCGATGCCAA
TGATACATTGCACATAACTGACAGGAAGTTATCCAAGTACTGTATAACATCTTGTTTATTATTTAATGTTTT
CTAAAAATAAAATGTGTAGTTTTTCCAAATGGCCTAATAAAAACAATATTGTAATAAAAAACACTGTGTAGTAAT

FIGURE 110

MDFLLLALVLVSSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCGWARQSWGQCQPVCQP
RCKHGECIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPD
GSCSSALTCSMANCQYGCDVVKGQIRQCPSPLHLAPDGRTCVDVDECATGRASCPRFRQC
VNTFGSYICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEYQG
DGLTCVYIPKVMIEPSPGIHVPKNGTILKGDGTGNNNWIIPDVGSTWPPKTPYIPPIITNRP
TSKPTTRPTPKPTPIPTPPPPPLPTELRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDN
RVQTDPOKPRGDVFSVLVHSCNFDHGLCGWIREKDNLDLHWEPIRDPAGGQYLTVSAAKAPGG
KAARLVLPFLGRLMHSGLCLSRHKVVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQI
TLRGADIKSESQR

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 273-277

Casein kinase II phosphorylation site.

amino acids 166-170, 345-349

Tyrosine kinase phosphorylation site.

amino acids 199-206

N-myristoylation site.

amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242,
421-427, 433-439, 462-468, 476-482

Aspartic acid and asparagine hydroxylation site.

amino acids 104-116, 186-198, 231-243

Cell attachment sequence.

amino acids 382-385

EGF-like domain cysteine pattern signature.

amino acids 75-87

FIGURE 112

MWLKVFTTFLSFATGACSGLKVTVPSTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERPH
TMPKYLLGSVNKSVVDPLEYQHKFTMPPNASLLINPLQFPDEGNYIVKVNIOGNGTSLASQ
KIQVTVDPPVTKPVVQIHPPSGAVEYVGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFS
PQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPIIYYGPGYGLQVNSDKGLKVGVEVFTV
DLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMDYVCCAYNNIT
GRQDETHFTVIIITSVGLEKLAQKGKSLSPASITGISLFLIISMCLLFLWKYQPYKVIKQK
LEGRPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRVSPASDCVSGQDLHSTV
YEVIQHIPAQQQDHPE

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 341-359

N-glycosylation site.

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208,
276-280, 308-312

Casein kinase II phosphorylation site.

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

Tyrosine kinase phosphorylation site.

amino acids 272-280

N-myristoylation site.

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237,
239-245

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 113

GCAAGCGGCGAAATGCGGCCCTCCGGGAGTCTTGCAAGTCCCTTGGCAGTCCTGGTGTCTGTT
GCTTTGGGTGCTCCCTGGAGCGCACGGGCGGCGGAGCAACGTTTCGCGTCATCAGCGGACGAGA
ACTGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCGTGGTGCCTGCT
TGTCAAATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGAGAAAGATCTTGAGGTTAA
TATTGCGAAAGTAGATGTCACAGAGCAGCCAGGACTGAGTGGACGGTTTTATCATAACTGCTC
TTCCTACTATTTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAG
AAGGACTTCATAAACTTTATAAGTGATAAAGAGTGAAGAGTATTGAGCCCGTTTCATCATG
GTTTGGTCCAGGTTCTGTTCTGATGAGTAGTATGTCAGCACTCTTTCAGCTATCTATGTGGA
TCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACT
GTTTTTGCTTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGC
AGATTGCCTTTGTCTCTCAAAAAGGCGCAGACCACAGCCATACCCATACCTTCAAAAAAT
TATTATCAGAATCTGCACAACCTTTGAAAAAGTGGAGGAGGAACAAGAGGCGGATGAAGAA
GATGTTTCAGAAGAAGAAGCTGAAAGTAAAGAAGGAACAAACAAAGACTTTCACAGAATGC
CATAAGACAACGCCTCTCTGGGTCCATCATTGGCCAACAGATAAATCCTAGTTAAATTTTATAG
TTATCTTAATATTATGATTTTGATAAAAAACAGAAGATTGATCATTTTGTGTTGGTTTGAAGTG
AACTGTGACTTTTTTGAATATTGCAAGGTTTCAGTCTAGATTGTCAATTAATTGAAGAGTCTA
CATTGAGAACATAAAGCACTAGGTATACAAGTTTGAATATGATTTAAGCACAGTATGATG
GTTTAAATAGTTCTCTAATTTTGAAAAAATCGTGCCAAGCAATAAGATTTATGTATATTTGT
TTAATAAACCCTTTTCAAGTCTGAGTTTTGAAAAATTTACATTTCCCAAGTATTGCATTAT
TGAGGTATTTAAGAAGATTATTTTAGAGAAAAATTTTCTCATTGATATAATTTTCTCTGT
TTTCACTGTGTGAAAAAAGAGATATTTCCCATAAATGGGAAGTTTGCCCATTTGTCTCAAG
AAATGTGTATTTCAGTGACAAATTCGTGGTCTTTTAGAGGTATATTTCCAAAAATTTCTTGT
ATTTTAGGTATGCAACTAATAAAAACTACCTTACATTAATTAATTACAGTTTTCTACACA
TGGTAATACAGGATATGCTACTGATTTAGGAAGTTTTTAAGTTCATGGTATTCTCTTGATTC
CAACAAAGTTTGATTTTCTCTTGTATTTTCTTACTTACTATGGGTTACATTTTATTTTT
CAAAATGGATGATAATTTCTTGGAAACATTTTATGTTTTAGTAACAGTATTTTTTGT
GTTTCAAACCTGAAGTTTACTGAGAGATCCATCAAATTGAACAATCTGTTGTAATTTAAAAAT
TTGGCCACTTTTTTCAGATTTTACATCATTCTTGCTGAACCTTCAACTTGAAATGTTTTTTT
TTTCTTTTGGATGTGAAGGTGAACATTCCTGATTTTGTCTGATGTGAAAAAGCCTTGGTA
TTTTACATTTTGAAAAATCAAGAAGCTTAATATAAAGTTTGCACTTCACTCAGGAAAAAG
CATCTTCTGTATATGTCTTAAATGTATTTTGTCCTCATATACAGAAAGTCTCTAATTGAT
TTTACAGTCTGTAATGCTTGTATTTTAAATAATAACATTTTATATTTTTTAAAGACAA
ACTTCATATTATCCTGTGTTCTTTCTGACTGGTAATATTGTTGGGATTTCACAGGTAAAA
GTCAGTAGGATGGAACATTTTGTGATTTTACTCCTTAAAGAGCTAGAATACATAGTTTT
CACCTTAAAGAAAGGGGAAATCATAAATCAATGAATCAATCACTGACCATACGTAGTAGAC
AATTTCTGTAATGTCCCTTCTTCTAGGCTCTGTTGCTGTGTAATCCATTAGATTTACAG
TATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAATATTGTACCATT
AAAGAGTTTGGATGTGTAACTGTGATGCCTTAGAAAAATCCTAAGCACAAAATAAACCT
TTCTAACCACTTCATTAAAGCTGAAAAAATAAAAAAAAAA

FIGURE 114

MAPSGSLAVPLAVLVLLWGPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNL
QPEWESFAEWGEDLEVNIKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI
NFISDKEWKSIIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFAL
ATLFSGLLLGLCMIFVADCLCPSKRRRPQYPYPSPKKLLSESAQPLKKVEEEQEADEEDVSE
EEAESKEGTNKDFPQNAIRQRSLGPSLATDKS

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 182-201

Casein kinase II phosphorylation site.

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

Tyrosine kinase phosphorylation site.

amino acids 107-115

N-myristoylation site.

amino acids 20-26, 192-198

Amidation site.

amino acids 25-29

FIGURE 115

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTAACTAATTCAACAAACGGGACCCCTT
CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCCAGTGGGACAGGCGGATTGGAAGAGCGGG
AAGGTCTCTGGCCAGAGCAGTGTGACACTTCCCTCTGTGACCATGAAACTCTGGGTGTCTGC
ATTGCTGATGGCCTGGTTTGGTGTCTGTAGCTGTGTGCAGGCCGAATTCCTACCTCTATTG
GGCACATGACTGACTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC
CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGAC
TAGCAAGTCAGCTGCTGATGCTGAGGGCTACTGGCTCACCCTGTGAATGCCTACAACTGG
TGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCTGACAGGACTCAGCTGCA
GGTTTTATCGCCAACTCTCTGTGCAGCGCAGTTCTTCCCCACTGATGAGGACGAGATAGG
AGCTGCCAAAGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCAGGCACAAATTTCCA
GAGGGGAACCTCCAGGAACCAAGTACCAGGCAATGTGAGTGTGGATGACTGCTTTGGGATG
GGCCGCTCGGCCTACAATGAAGGGGACTATTATCATACGGTGTGTGGATGGAGCAGGTGCT
AAAGCAGCTTGATGCCGGGAGGAGGCCACCACAACCAAGTCAAGGTGCTGGACTACCTCA
GCTATGCTGTCTTCCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCTGCTC
TCCCTTGACCCAAGCCACGAACGAGCTGGAGGGAACTCGCGGTACTTTGAGCAGTTATTGGA
GGAAGAGAGAAAAACGTTAAACAAATCAGACAGAAGCTGAGCTAGCAACCCAGAAAGGCA
TCTATGAGAGGCTGTGGACTACCTGCTGAGAGGGATGTTTACGAGAGCCTCTGTCTGTGG
GAGGGTGTCAAACCTGACACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAA
CAGGGCCCCACAGCTGTCTATTGCCCTTCAAAGAGGAGGACGAGTGGGACAGCCCCGACA
TCGTGAGTACTACGATGTCTATGTCTGATGAGGAAATCAGAGGATCAAGGAGATCGCAAAA
CCTAAACTTGCACGAGCCACCGTTCTGTATCCCAGACAGGAGTCTCTACTGTCTGCCAGCTA
CCGGGTTTCCAAAAGCTCTGGCTAGAGGAAGATGATGACCTGTGTGGCCCCGAGTAAATC
GTCGGATGCAGCATATCAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAAT
TATGGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGG
CCTCAAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTATGTAGAAGCTG
GTGGTGCCACCGTCTTCCCTGATCTGGGGCTGCAATTTGGCCTAAGAAAGGATCAGCTGTG
TTCTGGTACAACCTCTTGGGAGCGGGGAAGGTGACTACCGAACAGACATGTGCTGCCCTG
TGTGCTTGTGGGCTGCAAGTGGGTCTCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT
TGAGACCTTGTGGATCAACAGAAGTTGACTGACATCCTTTTCTGTCTCTCCCTTCTCTGGTCT
CTTCAGCCCATGTCAACGTGACAGACACCTTTGTATGTTCTTTGTATGTTCTTATCAGGCT
GATTTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGCGGACTCCTGT
GTGACTGAAGTCCCAGCCCTTCATTACGCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA
AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCTTTGTACCTCAGGTGTT
TTAGGTGTGAGATGTTTCAGTGAACCAAGTTCGTATACCTTGTTTACATGTTTGTTTTTAT
GGCATTTCTATCTATTGTGGCTTTACCAAAAAATAAAATGTCCCTACCAGAAAAA

FIGURE 116

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA
NKMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALDVLQDSAAGFIANLSVQRQFFP
TDEDEIGAALKMLRLQDITYRLDPGTISRGELPGTKYQAMLSVDDCFGMGRSAYNEGDDYYHTV
LWMEQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLSLDPSSHERAGGNLR
YFEQLLEEREKTLTNQTEAELATPEGIYERPVLDYLPEDVYESLCRGEKVLTTPRRQKRLF
CRYHHGNRAPQLLIAPFKEEEDWDSPHIVRYDVMSDEEIERIKEIAKPKLARATVRDPKTG
VLTVASYRVSKSSWLEEDDDPVVARVNRMRQHITGLTVKTAELLQVANYGVGGQYEPHFDFS
RRPFDSSGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGEGDYR
TRHAACPVLVGCWKVSNKWFHERGQEFLRPCGSTEV

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 115-119, 264-268

Glycosaminoglycan attachment site.

amino acids 490-494

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 477-481

Casein kinase II phosphorylation site.

amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270,
346-350, 365-369, 385-389, 457-461, 530-534

Tyrosine kinase phosphorylation site.

amino acids 71-80, 489-496

N-myristoylation site.

amino acids 14-20, 131-137, 171-177, 446-452

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 8-19

Leucine zipper pattern.

amino acids 213-235

FIGURE 117

GCAGTATTGAGTTTACTTCCTCCTCTTTTGTAGTGAAGACAGACCATAATCCCAGTGTGAGTGAAATTGATTGT
TTCATTATATACCGTTTGGCTGGGGGTTAGTTCCGACACCTTCACAGTTGAAGAGCAGGCAGAAGGAGTTGTGA
AGACAGGACAATCTTCTGGGGATGCTGGTCTCGGAAGCCAGCGGCTTGTCTGTCTTTGGCTCATTGACCC
CAGGTTCTCTGGTTAAACCTGAAGCCCTACTCTGGCTGGTGGCCATCAATCCATTGATCTTGAGGCTGTGCC
CCTGGGGCACCCACTGGGAGGCTTACCCACCATGGACTGAGCTCCCTGTGGCTCTGCTGGCCGAGCGCTTC
CCCTCATCTTAGGGCTGTCTCTGGGCTGACGCTGAGCCTCTGCGGGTTCTCTGGATCCAGGGGGAGGGAGAAG
ATCCCTGTGTGAGGGCTGTAGGGGAGCGAGGAGGGCCACAGAATCCAGATTGAGAGCTCGGCTAGACCAAAGTG
ATGAAGACTTCAAACCCCGATTGTCCCTACTACAGGGACCCAAACAGCCCTACAAGAAGGTGCTCAGGACTC
GGTACATCCAGACAGAGCTGGGCTCCGCTGAGCGGTTGCTGTGGCTGTCTGACCTCCCGAGCTACACTGTCCA
CTTTGGCCGTGGCTGTGAACCGTACGGTGGCCCATCACTTCCCTCGGTACTCTACTTCACTGGGAGCGGGGG
CCCCGGCTCCAGCAGGATGAGGTGGTGTCTCATGGGGATGAGCGGCCCGCTGGCTCATGTAGAGACCTTGC
GCCACCTTCAACACACTTTGGGGCCGACTACGACTGGTCTTTCATCATGAGGATGACACATATGTGCAGGCC
CCCGCTGGCAGCCCTTGTGGCCACCTCAGCATCAACCAAGACCTGTACTTAGGGCCGGCAGAGGAGTTTCATTG
GCGCAGGCGAGCAGGCCCGGTACTGTGATGGGGCTTTGGCTACCTGTTGTACAGGAGTCTCCTGCTTCGTCTGC
GGCCACATCTGGATGGCTGCCGAGGAGACATTCTAGTGCCCTCTGACGAGTGGCTTGGACGCTGCCTCATTG
ACTCTCTGGGCTGGCTGTGTCTCAGACACCAGGGGCGAGCAGTATCGCTCATTTGAACTGGCCAAAATAGGG
ACCTCGAGAAGGAAGGAGCTGGCTTTCTGAGTGCCCTCGCCGTGCAACCTGTCTCGAAGTGACCTCATGT
ACCGGCTCCACAAACGCTTCAGCGCTCTGAGGTGGAGCGGCTTACAGTGAATAGAACAACTGCAGGCTCAGA
TCCGGAACCTGACGCTGCTGACCCCGGAAGGGGAGCGAGGCTGAGCTGGCCCTTGGGCTCCTGCTCCTTCA
CACCACACTCTCGCTTTGAGGTGCTGGGCTGGGACTACTTCACAGAGCAGCACACTTCTCTGTGAGATGGG
CTCCCAAGTGCCCATTCAGGGGGCTAGCGGGCGGACGTGGTGATGCGTTGGAGACTGCCTGGAGCAGCTCA
ATCGGCTGTATGCACTCCCGCTGGCTTCCAGAAGCAGCGACTGTCTCAACGGCTATCGGCGCTTCGACCCAGCAC
GGGCGATGGAGTACCCCTGGACCTGTCTGTGAATGTGTGACACAGCGTGGGCACCGCGGGCCCTGGCTCGCA
GGGTGAGCTGTGCTGGCCACTGAGCCGGGTGGAATCCTACTATGCCCTATGTCACTGAGGCCACCCGAGTGC
AGCTGTGTGCTGCCACTCCTGGTGGCTGAAGCTGTGACGCCCCGCTTCTCTCGAGCGCTTTGAGCCAATGTCC
TGGAGCCAAGAGAATGCAATGTCTACCCCTGTGTGCTGTACGCGCCAAGAGAAGGTGGCCGTGGAGCTCAG
ACCCATTCTTTGGGTGAAGGCTGCAGCAGCGGAGTTAGAGCGACGCTACCTGGGACGAGGCTGGGCTGGCTCG
CTGTGCGAGCAGAGGCCCTTCCAGGTGCGACTCATGGACGTGCTCTCGAAGAAGCACCTGTGGGACACTCTCT
TCTTCTTACCAACCGTGTGGAAGAAGCTGGGCCGAAGTCTCAACCGCTGTGCGATGAATGCCATCTCTGGCT
GGCAGGCTCTTTCCAGTCCATTTCCAGGAGTTCAATCCTGCCCTGTCAACACAGAGATCACCCCAGGGCCCC
CGGGGGCTGGCCCTGACCCCCCTCCCTCCTGGTGTGACCCCTCCCGGGGGCTCTATAGGGGGAGATTG
ACCGCAGGCTTCTGCGAGGGCTGCTTCTACAACGCTGACTACCTGGCGGCCGAGCCCGCTGGCAGGTGAAC
TGGCAGGCGAGGAAGAGGAGGAGCCCTGGAGGGCTGGAGGTGATGATGTTTCTCCTCGGTCTCAGGGCTCC
ACCTCTTTTGGGCCGTGAGAGCCAGGCTGGTGCAGAAGTTCTCCCTGCGAGACTGACGCCACCGGCTCAGTGAAG
AATCTTACCACCGCTGCGGCTCAGCAACCTGGAGGGCTAGGGGGCCGTGCCCAGCTGGCTATGGCTCTCTTTG
AGCAGGAGCAGGCCAATAGCACTTAGCCCCCTGGGGCCCTAACTCATTAACCTTTCTTGTCTGCTCAGCC
CCAGGAAGGGCAAGCAGATGGTGGCAGATAGAGAATTGTGTCTGATTTTAAATATGAAATGTTATTAA
ACATGTCTTCTGCC

FIGURE 118

MRLSSLLALLRPALPLIILGLSLGCSLSLLRVSWIQEGGEDPCVEAVGERGGPQNPDSSRARLD
QSDDEDFKPRIVPYPYRDNPYKKVLRTRYIQTGLGSRRERLLVAVLTSRATLSTLAVAVNRVT
AHHFPRLLLYFTGQRGARAPAGMQVSVSHGDERPAWLMSETLRHLHTEFGADYDWFIMQDDTY
VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLSRLSLRLRLPHLDGCRG
DILSARPDDEWLGRCLIDSLGVGCVSQHQGQQYRSFELAKNRDPEKEGSSAFLSAFVHPVSE
GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLTVLTPGEAGLSWPVGLPAPFTPHSRFEV
LGWDYFTEQHTFSCADGAPKCP LQGASRADVGDALETALEQLNRRYQPRLRFQKQRLNGYR
RFDPPARGMEYTLDDLLECVTQRGHRRALARRVSLRLPLSRVEILPMPYVTEATRVQLVLPPL
VAEAAAAAPAFLEAFAANVLEPREHALLTLLL VYGPRGGRGAPDPFLGVKAAAAELERRYPG
TRLAWLAVRAEAPSQVRLMDVVSKKHPVDTLFFLT TVWTRPGPEVLNRCRMNAISGWQAFPP
VHFQEFNPALSPQRSPPGPPGAGDPDPPSPGADPSRGAPIGGRFDRQASAEGCFFYNADYLA
RARLAGELAGQEBEEALEGLEVMDFLRFSGLHLFRAVEPGLVQKFSRLRDCSPRLSEELYHR
CRLSNLEGLGGRAQLAMALFEQEQA NST

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 489-507

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389,
399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,
558-564, 651-657, 657-663, 672-678

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

FIGURE 119

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCTGTCCCCAAGCC
GTTCTAGACGCGGGAAAAATGCTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG
AAGCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA
TGCACCACCATGAGCATCATCACCTACAAGCTCCTAACAAAGAAGATATCTTGAAAAATTCA
GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTGAGTATACTGTATTATCCTTGTAAAACC
CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACTGTGACAAAGCAG
AGTTCTTCAGTTCTGAAAATGTTAAAGTGTGAGTCAATTAATATGGACACAAATGACATG
TGGTTAATGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACCTG
GTTCTTCTTGCACGCCCTACTACGTTTGCTATCATGAAAACCTAAAGTATTTTTTGTAA
AAAAGGATCCATCACAGCCTTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT
GTGGGTATGGAAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCT
CAATATCCCAGAAAAGTGTCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAAC
AGCTAGCAGTTTGCTGAAATATGCTGGAGTATTGCAGAAAATGCAGAAGATGCTGATGGA
AAAGATGTATTTAATACCAAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCC
CAACCAGGTAGTAGAAGGCTGTTGTTTCTGATATGGCTGTTACTTTAATGGACTGACTCCAA
ATCAGATGCATGTGATGATGATATGGGGTATACCGCCTTAGGGCATTTGGGCATATTTCAAT
GATGCATTGGTTTTCTTACCTCCAAATGGTTCTGACAATGACTGAGAAGTGGTAGAAAAGCG
TGAATATGATCTTTGTATAGGACGTGTGTTGTCTATTATTGTAGTAGTAACATACATATCCAA
TACAGCTGTATGTTTCTTTTCTTTTCTAATTGGTGGCACTGGTATAACCACACATTAAG
TCAGTAGTACATTTTAAATGAGGGTGGTTTTTTCTTTAAACACATGAACATTGTAAATG
TGTTGGAAGAAGTGTTTTAAAGAATAATAATTTGCAAATAAACTATTAATAAATATTATAT
GTGATAAATCTAAATTATGAACATTAGAAATCTGTGGGGCACATATTTTGTCTGATTGGTT
AAAAAATTTTAAACAGGCTTTAGCGTTCTAAGATATGCAAAATGATATCTCTAGTTGTGAATT
TGTGATTAAAGTAAACTTTTAGCTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCT
AAGCCTCCCCAAGTTCCAATGGATTGCGCTTCTCAAATGTACAACCTAAGCAACTAAAGAAA
ATTAAAGTGAAGTTGAAAAAT

FIGURE 120

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERME
LSKSFVRVYCIILVKPKDVSLWAAVKETWTKHCDKAEFFSSENVKVFESINMDTNDMWLMMRK
AYKYAFDKYRDQYNWFFLARPTTFAI IENLKYFLKKDPSQPFFYLGHITIKSGDLEYVGMIEGG
IVLSVESMKRLNSLLNIPEKCPQGGMWIKISEDKQLAVCLKYAGVFAENAEDADGKDVFNFT
KSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMYGVYRLRAFGHIFNDALVFL
PPNGSDND

Signal sequence:

amino acids 1-33

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389,
399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,
558-564, 651-657, 657-663, 672-672

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

THE UNIVERSITY OF CHICAGO

[illegible]

FIGURE 122

MNSSKSSETQCTERGCFSSQMFLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPEN
FTELSYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALS LKNCSAMGAHLVVINSQEEQ
EFLSYKKPKMREFFIGLSDQVVEGQWVDGTPLTKSLSFWDVGEFNNIATLED CATMRDSS
NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

Signal sequence:

amino acids 1-42

N-glycosylation site.

amino acids 2-6, 62-66, 107-111

Casein kinase II phosphorylation site.

amino acids 51-55, 120-124, 163-167, 175-179, 181-185

N-myristoylation site.

amino acids 15-21, 74-80, 155-161

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 27-38

FIGURE 123

GGGACTACAAGCGCGCGCGCTGCCGCTGGCCCCCTCAGCAACCCCTCGACATGCGCGCTGAGGCGGGCCACCGCGAC
 TC CGGCTCTGCGCTCGGCTGCGCTGACTTTCTTCTGCTGCTGCTTTTCAGGGGCTGCGCTGATAGGGGCTGTAAATC
 TCAAAATCCAGCAATCGAACCCAGTGGTACAGGAATTTGAAAGTGTGGAACCTGCTTGGCATCTACCGGATTCGC
 AGACAAGTGACCCCGAGGATCGAGTGGGAAGAAATTCAGATGAAACAAACACACATATGTGTTTTTGGACAACAAAA
 CTGTGCGAGTGAAGCCAGTGCACCCCTGTCTGTAGAGTCCGGAAGGCTGTACAGTAGGCGAAGATGGCAACATCTGC
 ACTCAGGCCCTTTATCGCTGTAGGTCGTGCTGCTGCAAAATGACCGCAAGGAAATGTATGAGATTGTGATCGAGTTAA
 CTGTGCGAGTGAAGCCAGTGCACCCCTGTCTGTAGAGTCCGGAAGGCTGTACAGTAGGCGAAGATGGCAACATCTGC
 ACTGCGCAGGAGGCTGAGGGGCCACCCCGGCGCTCAGTACAGCTGGTATCGCAATGATGTACCACTGCCACCGGATT
 CCAGAGGCCAATCCAGATTTCGCCAATTTCTTCTTCCACTTAAACTCTGAAACAGGCACTTTGGTGTTCACTGCTG
 TTCAAGAGGAGCTCTGCGGAGTACTACTGCAATGCTTCCAATGACCGAGGCTCAGCCAGGTTGTAGGAGGAGG
 AGATGTGAAGTCTATGACCTTCCCAATTTGGCGGAATTTATGGGGGGTCTGGTGTCTGCTTGTACTGCGCCCTGA
 TCAGTGTGGGCTATGCTGTGCATACGAGCGTGGCTACTTTCATCAACAAATAACAGGATGGAGAAAGTTACAAGA
 ACCCAGGGAACAGATGGAGTTAACTACATCCGCACTGACGAGGAGGGCGACTTCAGACACAAGTCATCGTTTG
 TGATCTGAGACCCGCGGTGTGGCTGAGAGCGCACAGAGCGCAGTGCACATACCTCTGCTAGAACTCCCTGTCAA
 GCGAGCGAGAGCTGATGCATCGACAGAGCTAGACACTCATTAGAAGCTTTTCGTTTGGCCAAAGTTGACCA
 CTACTCTTCTTACTCTAAACAGCCACATGAATAGAAGAAATTTCTCCTAAGATGGACCCGGTAAATATAACCA
 GGAAGCGAACTGGGTGCGTTCACTGAGTTGGGTCTTAATCTGTTTCTGGCCCTGATCCCGCATGAGTATTAGG
 GTGATCTTAAAGAGTTTCTCAGCTTAAACGCCCGTGTGGGCCCTGTGAAGCCAGCATGTTCCACCATGGTGTG
 CAGCAGCCACGACAGCACCATTGTGAGATGGCGAGGTGGCTGGACAGCAGCAGCGCATCTTCTTAAAGGCTCTGC
 GAAAAGGCTTCTTACACAGCAGCCCTTACTTTCATCGGCCACAGACACACCGCATCTTCTTAAAGGCTCTGC
 TGATCGGTGTGTCAGTGTCCATTTGGGAGAAAGCTTTTGGATCAGCATTTTGTAAACAAACCAAAATCAGGA
 GTAAATTTGGTGTGCGAAGAGGATCTTGCTGAGGAACCTGCTTCCACAGGCTCTCAGGATTTAAGGAAA
 ACCTTGTCTTAGGCTAAGTCTGAAATGGTACTGAAATATGCTTTCTATGGGTCTGTGTTATTTTATAAAATTT
 TACATCTAAATTTTGTCAAGGATGTATTTTGATTAATGAAAAGAAATTTCTATTTAACTGTAAATATATTTGT
 CATACAATGTAAATAACCTATTTTTTAAAAAAGTCCAACCTAAGGTAGAAGTTCCAAGCTACTAGTGTAAAT
 TGGAAAATATCAATAATTAAAGATTTTACCCAAGGAATCTCTCATGTGGAAGTTACTGTGATGTTCCTTTCT
 CACACAAGTTTTAGCCTTTTTCACAAAGGAACTCATATCTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT
 TAAAAATTCAGTTAAAGCAATTTGAAATCAGTTTGCATCTCTTCAAAAGAAACCTCTCAGGTTAGCTTTGAAT
 GCCTCTTCTGTAGATGACTAGGACAGTCTGTACCAGAGGGCCACCCAGAAAGCCCTCAGATGTACATACAGATG
 CCAGTCAGCTCTCGGGTTGGCCAGGCGCCCGGCTTAGCTCTACTGTGCTCGCTGTGTCGCGAGGAGGCCCT
 GCCATCTTGGGCCCTGGCAGTGGCTGTGTCCAGTGTGCTTACTCAGCTGGGCCCTGCTTCATCCAGCACAGC
 TCTCAGGTGGGCACTGCAGGGAACCTGGTGTCTTCCATGTAGCGTCCAGCTTTGGGCTCTGTAAACAGACCTCT
 TTTTGGTTATGGATGGCTCAAAAATAGGGCCCCCAATGCTATTTTTTTTTTAAAGTTGTTTAAATTTATTTGTT
 AAGATTGTCTAAGGCCAAAGGCAATTCGGAATCAAGTCTGTCAAGTACAATAACATTTTAAAAAGAAATGGAT
 CCCATGTTCTCTTTGCCACAGAGAAAGCACCAGAGCCACAGGCTCTGTGCAATTTCAAAACAAACCATGAT
 GAGGTGGCGGCGAGTCCAGCTTTTAAAGAACGTCAAGTGGAGCAGCCAGGTGAAAGGCTGGCGGGGAGGAAAG
 TGAAACGCCGTAATCAAAGCAGTTTCTAATTTTGAATTTTAAATTTTCAATCGCGGGAGACACTGCTCCCATT
 TGTGGGGGACATTAGACAACATCACTCAGAAGCTGTGTTCTTCAAGAGCAGTGTCTCAGCCCTCACA TGCCCT
 GCGCTCTGCGACTCAGGACTGAAGTGTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCTGGA
 GAATGGGCTCTCACTACTTGTCTTTCAGCTTCCAGTGTCTGGGTTTTTATATCTTGACAGCTTTTTTTTT
 AATTGCATACATGAGACTGTGTTGACTTTTTTTAGTTATGTGAAACACTTTGCGCAGGCGCCCTGGCAGAGGCA
 GGAAGTGCCTCAGCAGTGGCTCAGTGTCTTCCGTGCTGTGCTCAGTGCCTGGATGCTAGCATGCAAGTTTCT
 CCTCCATCATGCCACTTGGTAGAGAGGATGGCTCCCCCCTCAGCGTTGGGGATTACAGCTCAGGCTCCTCT
 TCTTGTGTTGTATAGTGTATAGGAGTACCTTATGTGCCCCCTCTTATACCTTAAACCTTCAACATGTAGTGA
 TGGGAAACAGGCTCTGAAAGTAGAGAGAAGTGAAGTAGAGTCTGGGAAGTAGCTGCTTAACTAGACTAGACTAGA
 CGGAAAAGGAATACTCTGTGTATTTTGAAGATATGAATGTGATCAAGACTCGAGCGGATCAGGCGCTGATATT
 GCCCTTGGATGATGTGCTGTACACAGATGCTACAGACTGTACTAACACACCGTAAATTTGGCATTGTGTTAAC
 CTCATTTATAAAGCTTCAAAAAACCA

FIGURE 124

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77624
><subunit 1 of 1, 310 aa, 1 stop
><MW: 35020, pI: 7.90, NX(S/T): 3
MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVQQEFESVELSCIITDSQTS
PRIEWKKIQDEQTTTVFFDNKIQGDLAGRAEILGKTSLSKIWNVTRRDSALYRCEVVARNDRK
EIDEIVIELTVQVKPVTVPVCRVPKAVPVGKMATLHCQSESEGHPRPHYSWYRNDVPLPTDSRA
NPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEYVDLNIIGGIIG
VLVVLAVLALITLGICCAYYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI
```

Important features of the protein:

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 243-263

N-glycosylation sites.

amino acids 104-107, 192-195

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 107-110

Casein kinase II phosphorylation site.

amino acids 106-109, 296-299

Tyrosine kinase phosphorylation site.

amino acids 69-77

N-myristoylation sites.

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267